

(19) World Intellectual Property
Organization
International Bureau



(43) International Publication Date
15 September 2005 (15.09.2005)

PCT

(10) International Publication Number
WO 2005/084430 A1

(51) International Patent Classification⁷: **A01K 67/027**

(US). LEAVITT, Markley, C. [US/US]; 1140 Latham Drive, Watkinsville, GA 30677 (US).

(21) International Application Number:
PCT/US2005/006072

(74) Agent: YESLAND, Kyle; Legal Department, AviGenics, Inc., 111 Riverbend Road, Athens, GA 30605 (US).

(22) International Filing Date: 28 February 2005 (28.02.2005)

(81) Designated States (unless otherwise indicated, for every kind of national protection available): AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BW, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NA, NI, NO, NZ, OM, PG, PH, PL, PT, RO, RU, SC, SD, SE, SG, SK, SL, SM, SY, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, YU, ZA, ZM, ZW.

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
10/790,455 1 March 2004 (01.03.2004) US
10/811,136 26 March 2004 (26.03.2004) US
10/940,315 14 September 2004 (14.09.2004) US

(71) Applicant (for all designated States except US): AVIGENICS, INC [US/US]; Legal Department, 111 Riverbend Road, Athens, GA 30605 (US).

(84) Designated States (unless otherwise indicated, for every kind of regional protection available): ARIPO (BW, GH, GM, KE, LS, MW, MZ, NA, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HU, IE, IS, IT, LT, LU, MC, NL, PL, PT, RO, SE, SI, SK, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

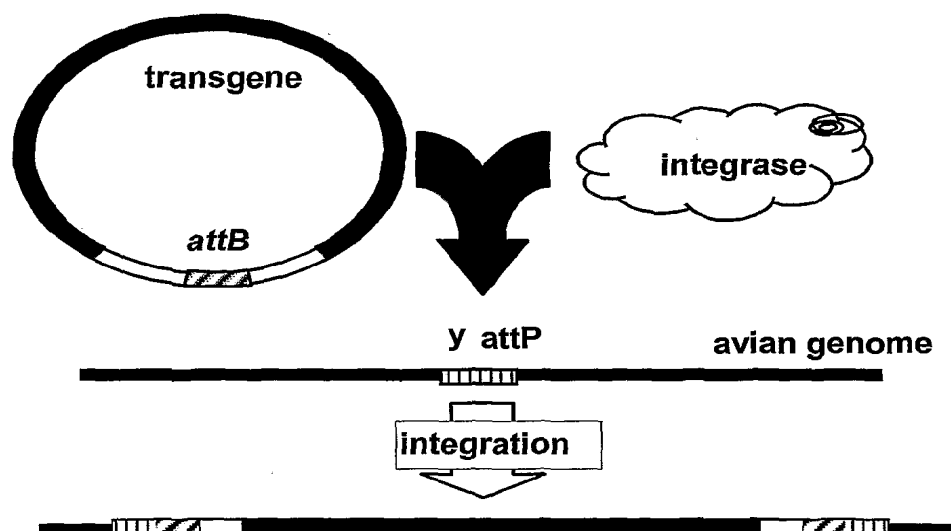
(72) Inventors; and

(75) Inventors/Applicants (for US only): CHRISTMANN, Leandro [BR/US]; 1311 Victoria Road, Watkinsville, GA 30677 (US). EBERHARDT, Dawn [US/US]; 202 Stella Henley Road, Danielsville, GA 30633 (US). HARVEY, Alex, J. [US/US]; 188 Sunset Drive, Athens, GA 30606

Published:
— with international search report

[Continued on next page]

(54) Title: GENERATION OF TRANSGENIC AVIANS



(57) Abstract: The invention includes transchromosomal avians and transchromosomal avian cells and methods for the introduction of artificial chromosomes into the genome of avians and avian cells.

WO 2005/084430 A1



— *before the expiration of the time limit for amending the claims and to be republished in the event of receipt of amendments*

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

GENERATION OF TRANSGENIC AVIANS

The present application claims priority to U.S. Patent Application
5 10/940,315, filed September 14, 2004; U.S. Patent Application No. 10/811,136,
filed March 26, 2004 and U.S. Patent Application No. 10/790,455, filed March 1,
2004.

Field of the Invention

10 The present invention relates to the field of biotechnology, and more
specifically to the field of genome modification. Disclosed herein are
compositions including chromosomes and vectors, and methods of use thereof, for
the generation of genetically transformed cells and animals.

Background

Transgenic technology to convert animals into "bioreactors" for the
production of specific proteins or other substances of pharmaceutical interest
(Gordon et al, 1987, Biotechnology 5: 1183-1187; Wilmut et al, 1990,
Theriogenology 33: 113-123) offers significant advantages over more conventional
20 methods of protein production by gene expression. For example, recombinant
nucleic acid molecules have been engineered and incorporated into transgenic
animals so that an expressed heterologous protein may be joined to a protein or
peptide that allows secretion of the transgenic expression product into milk or
urine, from which the protein may then be recovered.

25 Another system useful for heterologous protein production is the avian
reproductive system. The production of an avian egg begins with formation of a
large yolk in the ovary of the hen. The unfertilized oocyte or ovum is positioned
on top of the yolk sac. After ovulation the ovum passes into the infundibulum of
the oviduct where it is fertilized, if sperm are present, and then moves into the
30 magnum of the oviduct which is lined with tubular gland cells. These cells secrete
the egg-white proteins, including ovalbumin, lysozyme, ovomucoid, conalbumin
and ovomucin into the lumen of the magnum where they are deposited onto the
avian embryo and yolk. The hen oviduct offers outstanding potential as a protein

bioreactor because of the high levels of protein production, the promise of proper folding and post-translation modification of the target protein, the ease of product recovery, and the relatively short developmental period of chickens.

One method for creating permanent genomic modification of a eukaryotic cell is to integrate an introduced DNA into an existing chromosome. Retroviruses have so far proven to be the method of choice for efficient integration. However, retroviral integration is directed to a number of insertion sites within the recipient genome so that positional variation in heterologous gene expression can be evident. Unpredictability as to which insertion site is targeted introduces an undesirable lack of control over the procedure. An additional limitation of the use of retroviruses is that the size of the nucleic acid molecule encoding the virus and heterologous sequences may be limited to about 8 kb. In addition, retroviruses may include undesirable features such as splice sites. Although wild-type adeno-associated virus (AAV) often integrates at a specific region in the human genome, replication deficient vectors derived from AAV do not integrate site-specifically possibly due to the deletion of the toxic rep gene. In addition, homologous recombination produces site-specific integration, but the frequency of such integration usually is typically low.

An alternative method for delivering a heterologous nucleic acid into the genome is the use of a site-specific enzymes that can catalyze the insertion of nucleic acids into chromosomes. These enzymes recognize relatively short unique nucleic acid sequences that serve for both recognition and recombination. Examples include Cre (Sternberg & Hamilton, 1981, J. Mol. Biol. 150: 467-486, 1981), Flp (Broach et al, 1982, Cell 29: 227-234, 1982) and R (Matsuzaki et al, 1990, J. Bact. 172: 610-618, 1990).

A novel class of phage integrases that includes the integrase from the phage phiC31 can mediate highly efficient integration of transgenes in mammalian cells both in vitro and in vivo (Thyagarajan et al, Mol. Cell Biol. 21: 3926-3934, 2001). Constructs and methods of using recombinase to integrate heterologous DNA into a plant, insect or mammalian genome are described by Calos in U.S. Patent Serial No. 6,632,672, the disclosure of which is incorporated in its entirety herein by reference.

The phiC31 integrase is a member of a subclass of integrases, termed serine recombinases, that include, for example, R4 and TP901-1. Unlike the phage lambda integrases, which belong to a tyrosine class of recombinases, the serine integrases do not require cofactors such as integration host factor. The phiC31 integrase normally mediates integration of the phiC31 bacteriophage into the genome of *Streptomyces* via recombination between the attP recognition sequence of the phage genome and the attB recognition sequence within the bacterial genome. When a plasmid is equipped with a single attB site, phiC31 integrase will detect and mediate crossover between the attB site and a pseudo-attP site within the mammalian genome. Such pseudo-attP integration sites have now been identified in the mouse and human genomes. If the heterologous DNA is in a circular or supercoiled form, the entire plasmid becomes integrated with attL and attR arms flanking the nucleic acid insert. PhiC31 integrase is not able to mediate the integration into genomic DNA of sequences bearing attP sites.

Integration mediated by certain integrases, such as PhiC31 integrase-mediated integration, results in the destruction of the recognition or recombination sites themselves so that the integration reaction is irreversible. This will bypass the primary concern inherent with other recombinases, i.e., the reversibility of the integration reaction and excision of the inserted DNA.

Another method for the stable introduction of heterologous nucleic acid (e.g., large heterologous nucleic acids) into a genome is by the use of an artificial chromosome. Artificial chromosomes for expression of heterologous genes in yeast are available, but artificial chromosomes being delivered to avians has not previously been achieved.

Therefore, it is an object of the invention to produce transgenic animals with large nucleic acid segments integrated into their genome and to provide avians which include an artificial chromosome in their genome.

Summary of the Invention

Integration of a transgene into a defined chromosomal site is useful to improve the predictability of expression of the transgene, which is particularly advantageous when creating transgenic vertebrate animals such as, transgenic

avians. Transgenesis by methods that randomly insert a transgene into a genome are often inefficient since the transgene may not be expressed at the desired levels or in desired tissues.

5 The present invention relates to methods of modifying the genome of vertebrate cells (e.g., production of transgenic vertebrates) and to such cells with modified genomes and their progeny. In one embodiment, the methods provide for introducing into vertebrate cells a first recombination site such that the recombination site is inserted into the vertebrate cell genome. Typically, in such
10 embodiments, the genome does not normally include this first recombination site prior to the recombination site introduction. Methods of the invention may also include introducing a nucleotide sequence comprising a second recombination site and a sequence of interest such as a coding sequence into the vertebrate cell or progeny of the vertebrate cell. The nucleotide sequence comprising the second recombination site and the sequence of interest such as a coding sequence may be
15 introduced into the vertebrate cell before, at about the same time as or after the introduction of the first recombination site. Additionally, the present methods may include introducing into the vertebrate cell or progeny cell thereof a substance which facilitates insertion of the nucleotide sequence comprising the second recombination site and the sequence of interest proximal to the first recombination
20 site. For example, the nucleotide sequence comprising the second recombination site and the sequence of interest may be inserted adjacent to or internally in the first recombination site. In one very useful embodiment, the first recombination site and/or the nucleotide sequence comprising the second recombination site and the sequence of interest are stably incorporated into the genome of the cell.

25 The present invention contemplates the genomic modification of any useful vertebrate cells including, but not limited to, avian cells. Examples of cells which may have their genomes modified in accordance with the present invention include, without limitation, reproductive cells including sperm, ova and embryo cells and nonreproductive cells such as tubular gland cells.

30 The present invention also relates to methods of producing transgenic vertebrate animals and to the transgenic animals produced by the methods and to their transgenic progeny or descendants. The invention also includes the transgenic

cells included in or produced by the transgenic vertebrate animals. Examples of such cells include, without limitation, germ line cells, ova, sperm cells and protein producing cells such as tubular gland cells. In one useful embodiment, the transgenic vertebrate animals of the invention are transgenic avians. Transgenic
5 avians of the invention may include, without limitation, chickens, turkeys, ducks, geese, quail, pheasants, parrots, finches, hawks, crows or ratites including ostrich, emu or cassowary.

In accordance with the present invention, methods of producing transgenic vertebrate animals can include introducing into an embryo of a vertebrate animal a
10 first recombination site such that the recombination site is present in sperm or ova of a mature vertebrate animal developed from the embryo. In one useful embodiment, the embryo does not normally include the first recombination site in its genome prior to the recombination site introduction. The methods may also include introducing a nucleotide sequence comprising a second recombination site
15 and a sequence of interest such as a coding sequence into the embryo of the vertebrate animal. The first recombination site and/or the nucleotide sequence comprising the second recombination site and a sequence of interest may be introduced into the embryo of the vertebrate animal before the embryo is fertilized (i.e., when an ovum), at about the same time as introduction of the sperm into the
20 ovum or after fertilization.

The methods can also include introducing the nucleotide sequence comprising a second recombination site and a sequence of interest into an ovum or a sperm of a mature vertebrate animal developed from the embryo (or its descendents) into which the first recombination site was introduced. In one
25 embodiment, the nucleotide sequence comprising a second recombination site and a sequence of interest is introduced into the ovum from the mature vertebrate animal before the ovum is fertilized. In another embodiment, the nucleotide sequence comprising a second recombination site and a sequence of interest is introduced into the ovum at about the time of fertilization. In one particularly
30 useful embodiment, the nucleotide sequence comprising a second recombination site and a sequence of interest is introduced into the ovum after the ovum is fertilized (when an embryo).

The methods may include, upon addition of the nucleotide sequence comprising a second recombination site and a sequence of interest to an embryo, ovum or sperm, introducing into the embryo, ovum or sperm, a substance which facilitates insertion of the nucleotide sequence comprising the second
5 recombination site and the sequence of interest proximal to the first recombination site. For example, the nucleotide sequence comprising the second recombination site and the sequence of interest may be inserted adjacent to or internally in the first recombination site. In one useful embodiment, the methods include introducing into an embryo comprising the first recombination site in its genome, a substance
10 which facilitates insertion of the nucleotide sequence comprising the second recombination site and the sequence of interest proximal to the first recombination site.

In one useful embodiment, these methods include fertilizing an ovum with sperm comprising the first recombination site. The methods can include also
15 introducing into the ovum a nucleotide sequence comprising a second recombination site and a sequence of interest such as a coding sequence and a substance which facilitates insertion of the nucleotide sequence comprising the second recombination site and sequence of interest proximal to (e.g., adjacent to or internally in) the first recombination site. It is contemplated that the nucleotide
20 sequence comprising a second recombination site and a sequence of interest may be introduced into the ovum before or after fertilization by the sperm or at about the same time as fertilization.

In one very useful embodiment of the methods disclosed herein, the nucleotide sequence comprising the second recombination site and the sequence of
25 interest is stably incorporated into the genome of the embryo, ovum or sperm.

The methods disclosed herein typically eventually include exposing a fertilized ovum to conditions which lead to the development of a viable transgenic vertebrate animal.

In one embodiment, the nucleotide sequence of interest includes an
30 expression cassette. Optionally, the nucleotide sequence of interest may include a marker such as, but not limited to, a puromycin resistance gene, a luciferase gene, EGFP-encoding gene, and the like.

Typically, in accordance with methods known in the art or methods disclosed herein, the embryo of the vertebrate animal or fertilized ovum of a mature vertebrate animal of the invention is exposed to conditions which lead to the development of a viable transgenic vertebrate animal.

5 Embryos that are useful in the present methods include, without limitation, stage I, stage II, stage III, stage IV, stage V, stage VI, stage VII, stage VIII, stage IX, stage X, stage XI and stage XII embryos.

In one embodiment, the nucleotide sequence included with the second recombination site of interest is a coding sequence. The nucleotide sequence of
10 interest included with the second recombination site can be of any useful size. For example, and without limitation, the nucleotide sequence of interest may be from about 0.1 kb to about 10 mb, for example, about 1 kb to about 1 mb. In one embodiment, the nucleotide sequence of interest is about 5 kb to about 5 mb in size, for example, about 5 kb to about 2 mb, e.g., about 8 kb to about 1 mb. In one
15 embodiment, the nucleotide sequence of interest is about 0.5 kb to about 500 kb.

The first recombination site and/or the nucleotide sequence which includes the second recombination site and a sequence of interest such as a coding sequence may be introduced into cells, embryos (i.e., fertilized ova) or sperm by any useful method. These useful methods include, without limitation, cell fusion, lipofection,
20 transfection, microinjection, calcium phosphate co-precipitation, electroporation, protoplast fusion, particle bombardment and the like. In addition, the first recombination site or nucleotide sequence comprising the second recombination site and the sequence of interest may be introduced into cells, embryos, ova or sperm in the presence of a cationic polymer such as PEI and/or other substances
25 disclosed elsewhere herein or known in the art.

In one embodiment, recombination sites employed in the present invention are isolated from bacteriophage and/or bacteria. For example, the recombination sites may be attP sites or attB sites.

The substance which facilitates insertion of the second recombination site
30 and a sequence of interest may be an enzyme. In one embodiment, the substance is a site specific recombinase. In one useful embodiment, the substance which facilitates insertion of the nucleotide sequence is nucleic acid, for example, DNA

or RNA. The DNA or RNA may include modified nucleosides as described elsewhere herein or are known to those of skill in the art. In one embodiment, modified nucleosides are employed to extend the half-life of RNA or DNA molecules employed in the present invention. For example, it may be desirable to
5 extend the half life of the RNA or DNA molecules in the presence of a cellular environment. In one useful embodiment, the nucleic acid encodes an enzyme such as a site specific recombinase.

Nonlimiting examples of site specific recombinases which may be employed herein either as protein or encoded by nucleic acid include serine
10 recombinases and tyrosine recombinases. Examples of serine recombinases which may be employed include, without limitation, EcoYBCK, Φ C31, SCH10.38c, SCC88.14, SC8F4.15c, SCD12A.23, Bxb1, WwK, Sau CcrB, Bsu CisB, TP901-1, Φ 370.1, Φ 105, Φ FC1, A118, Cac1956, Cac1951, Sau CcrA, Spn, TnpX, TndX, SPBc2, SC3C8.24, SC2E1.37, SCD78.04c, R4, Φ Rv1, Y4bA and Bja serine
15 recombinases.

In one embodiment of the invention, the present methods include introducing an integration host factor into a cell (e.g., an embryo) to facilitate genomic integration. Such integration host factors may be particularly useful when employing certain substances such as tyrosine recombinases as disclosed herein.

20 The nucleotide sequence of interest may include a coding sequence. The coding sequence may encode any useful protein. In one useful embodiment, the sequence of interest encodes a pharmaceutical or therapeutic substance. The invention contemplates the production of any useful protein based pharmaceutical or therapeutic substance. Examples of pharmaceutical or therapeutic substances
25 include without limitation at least one of a light chain or a heavy chain of an antibody (e.g., a human antibody) or a cytokine. In one embodiment, the pharmaceutical or therapeutic composition is interferon, erythropoietin, or granulocyte-colony stimulating factor. In one embodiment, the transgenic animal is an avian and the sequence of interest encodes a polypeptide present in eggs
30 produced by the avian.

In one embodiment, integrases such as phage integrases, for example, serine recombinases, such as the integrase from phage ϕ C31, can mediate the efficient

integration of transgenes into target cells both in vitro and in vivo. In one embodiment, when a plasmid is equipped with a single attB site, the integrase detects attP homologous sequences, termed pseudo-attP sites, in a target genome and mediates crossover between the attB site and a pseudo attP site.

5 In one embodiment, once delivered to a recipient cell, for example, an avian cell, the phiC31 integrase mediates recombination between the att site within the nucleic acid molecule and a bacteriophage attachment site within the genomic DNA of the cell. Both att sites are disrupted and the nucleic acid molecule, with partial att sequences at each end, is stably integrated into the genome attP site. The
10 phiC31 integrase, by disrupting the att sites of the incoming nucleic acid and of the recipient site within the cell genome can preclude any subsequent reverse recombination event that would excise the integrated nucleic acid and reduce the overall efficiency of stable incorporation of the heterologous nucleic acid.

 Following delivery of the nucleic acid molecule and a source of integrase
15 activity into a cell population and integrase-mediated recombination, the cells may be returned to an embryo. In the case of avians, late stage blastodermal cells may be returned to a hard shell egg, which is resealed for incubation until hatching. Stage I embryos may be directly microinjected with the polynucleotide and source of integrase activity, isolated, transfected and returned to a stage I embryo which is
20 reimplanted into a hen for further development. Additionally, the transfected cells may be maintained in culture in vitro.

 The present invention provides novel methods and recombinant polynucleotide molecules for transfecting and integrating a heterologous nucleic acid molecule into the genome of a cell of a vertebrate animal, such as an avian.
25 Certain methods of the invention provide for the delivery to a cell population a first nucleic acid molecule that comprises a region encoding a recombination site, such as a bacterial recombination site or a bacteriophage recombination site. In one embodiment, a source of integrase activity is also delivered to the cell and can be in the form of an integrase-encoding nucleic acid sequence and its associated
30 promoter or as a region of a second nucleic acid molecule that may be co-delivered with the polynucleotide molecule. Alternatively, integrase protein itself can be delivered directly to the target cell.

The recombinant nucleic acid molecules of the present invention may further comprise a heterologous nucleotide sequence operably linked to a promoter so that the heterologous nucleotide sequence, when integrated into the genomic DNA of a recipient cell, can be expressed to yield a desired polypeptide. The nucleic acid molecule may also include a second transcription initiation site, such as an internal ribosome entry site (IRES), operably linked to a second heterologous polypeptide-encoding region desired to be expressed with the first polypeptide in the same cell.

The present invention provides modified isolated artificial chromosomes useful as vectors to shuttle transgenes or gene clusters into a genome of an avian. By delivery of the modified chromosome to a recipient cell, the target cell, and progeny thereof, become trisomic or transchromosomic. The additional chromosome will typically not affect the subsequent development of the recipient cell and/or embryo, nor interfere with the reproductive capacity of an adult bird developed from such cells or embryos. The chromosome will also be stable within the genome of the cells of the adult bird or within isolated avian cells. The invention provides methods to isolate a population of chromosomes for delivery into embryos or early cells of avians, for example, chickens.

The methods can include inserting a lac-operator sequence into an isolated chromosome and, optionally, inserting a desired transgene sequence within the same chromosome. The lac operator region is typically a concatamer of a plurality of lac operators for the binding of multiple lac repressor molecules. A recombinant DNA molecule is constructed that includes an identified region of the target chromosome, a recombination site such as attB or attP, and the lac-operator concatamer. The recombinant molecule is delivered to an avian cell, and homologous recombination will integrate the heterologous polynucleotide and the lac-operator concatamer into the targeted chromosome. A tag-polypeptide, such as the GPF-lac-repressor fusion protein, binds to the lac-operator sequence for identification and isolation of the genetically modified chromosome. The tagged mitotic chromosome can be isolated using, for instance, flow cytometry.

Among other things, the present invention relates to transchromosomic avians. In a particular aspect, the invention provides for G0 transchromosomic

avians (e.g., germline chimeric transchromosomic avians) which can produce germline transchromosomic offspring (e.g., G1 and G2 germline transchromosomic offspring).

5 Examples of avians which are contemplated for use herein include, without limitation, chicken, turkey, duck, goose, quail, pheasants, parrots, finches, hawks, crows and ratites including ostrich, emu and cassowary.

In one useful aspect, the artificial chromosome employed herein includes a centromere. Any useful centromere may be employed in the present invention including, without limitation, centromeres from insects, mammals or avians.

10 In one particularly useful embodiment, the artificial chromosomes used herein include a heterologous nucleotide sequence. The nucleotide sequence may be heterologous to the avian and/or heterologous to the artificial chromosome. In one useful embodiment, the heterologous nucleotide sequence includes a coding sequence for a therapeutic substance. In addition, the heterologous nucleotide
15 sequence may include a gene expression controlling region. Any useful gene expression controlling region may be employed in the invention. For example, and without limitation, the gene expression controlling region may include a lysozyme promoter, an ovomucin promoter, a conalbumin promoter, an ovomucoid promoter and/or an ovalbumin promoter or functional portions thereof. See, for example,
20 US Patent Application No. 10/114,739, filed April 1, 2002; US Patent Application No. 10/856,218, filed May 28, 2004 and US Patent Application No. 10/733,042, filed December 11, 2003. The disclosure of each of these patent applications is incorporated herein by reference in its entirety. In one useful embodiment, the product of the heterologous nucleotide sequence (e.g., therapeutic substance) is
25 delivered to the avian egg (e.g., the egg white) during production of the egg in the avian. The invention also includes the eggs produced by the avians produced by these methods and other methods disclosed herein.

One useful aspect of the invention relates to methods of producing transchromosomic avians. In one embodiment, the methods include substantially
30 purifying a chromosome followed by introducing the purified chromosome into an avian embryo and thereafter maintaining the embryo under conditions suitable for the embryo to develop and hatch as a chick. In one embodiment, the methods

include inserting a heterologous nucleotide sequence into the chromosome before substantially purifying the chromosome. In one embodiment, the chromosome is introduced into the avian embryo by microinjection; however, any useful method to introduce the chromosome into the avian embryo is within the scope of the present invention.

It is contemplated that the chromosome may be introduced into the embryo by delivering the chromosome to an avian cell before or after fertilization. For example, the chromosome may be introduced into an ovum or a sperm before fertilization. In another example, the chromosome is introduced into a cell of an embryo (e.g., stage I to stage XII embryo). In one embodiment, the chromosome is introduced into an early stage embryo, for example, and without limitation, a stage I embryo. In one embodiment, the chromosome is introduced into a germinal disc.

The methods provide for the introduction of any useful number of chromosomes into the avian embryo in order to produce a transchromosomal avian. For example, and without limitation, between 1 and about 10,000 chromosomes may be introduced into the embryo. In another example, between 1 and about 1,000 chromosomes may be introduced into the embryo.

The invention also provides for transchromosomal avian cells wherein the artificial chromosome includes a nucleotide sequence which encodes a therapeutic substance. The cells may be isolated from transchromosomal avians and thereafter grown in culture. The invention also contemplates the production of the transchromosomic avian cells by stable introduction of the artificial chromosome into cultured avian cells. Any useful method may be employed for the introduction of the artificial chromosome into the cultured cells including, without limitation, lipofection or microinjection.

Another aspect of the present invention is a cell, for example, an avian cell, genetically modified with a transgene vector by the methods of the invention. For example, in one embodiment, the transformed cell can be a chicken early stage blastodermal cell or a genetically transformed cell line, including a sustainable cell line. The transfected cell may comprise a transgene stably integrated into the nuclear genome of the recipient cell, thereby replicating with the cell so that each progeny cell receives a copy of the transfected nucleic acid. One useful cell line

for the delivery and integration of a transgene comprises a heterologous attP site that can increase the efficiency of integration of a polynucleotide by an integrase, such as phiC31 integrase and, optionally, a region for expressing the integrase.

Another aspect of the present invention is methods of expressing a heterologous polypeptide in a cell by stably transfecting a cell by using site-specific
5 integrase-mediation and a recombinant nucleic acid molecule, as described above, and culturing the transfected cell under conditions suitable for expression of the heterologous polypeptide under the control of a transcriptional regulatory region.

Yet another aspect of the present invention concerns transgenic vertebrate
10 animals, such as birds, for example chickens, comprising a recombinant nucleic acid molecule and which may (though optionally) express a heterologous gene in one or more cells in the animal. For example, in the case of avians, embodiments of the methods for the production of a heterologous polypeptide by the avian tissue involve providing a suitable vector and introducing the vector into embryonic
15 blastodermal cells containing an attP site together with an integrase, for example, a serine recombinase such as phiC31 integrase, so that the vector can integrate into the avian genome at the attP site which has been engineered into the cell genome. A subsequent step may involve deriving a mature transgenic avian from the transgenic blastodermal cells by transferring the transgenic blastodermal cells to an
20 embryo, such as a stage X embryo (e.g., an irradiated stage X embryo), and allowing that embryo to develop fully, so that the cells become incorporated into the bird as the embryo is allowed to develop. In one embodiment, sperm from a G0 bird positive for the transgene is used to inseminate a chicken giving rise to a fully transgenic G1 generation.

25 One approach may be to transfer a transfected nucleus to an enucleated recipient cell which may then develop into a zygote and ultimately an adult animal. The resulting animal is then grown to maturity.

In the transgenic vertebrate of the present invention, the expression of the transgene may be restricted to specific subsets of cells, tissues or developmental
30 stages utilizing, for example, trans-acting factors acting on the transcriptional regulatory region operably linked to the polypeptide-encoding region of interest of the present invention and which control gene expression in the desired pattern.

Tissue-specific regulatory sequences and conditional regulatory sequences can be used to control expression of the transgene in certain spatial patterns. Moreover, temporal patterns of expression can be provided by, for example, conditional recombination systems or prokaryotic transcriptional regulatory sequences. By
5 inserting an integration site such as attP into the genome, it is believed that expression of an integrated coding sequence will be much more predictable.

The invention can be used to express, in large yields and at low cost, a wide range of desired proteins including those used as human and animal pharmaceuticals, diagnostics, and livestock feed additives. Proteins such as growth
10 hormones, cytokines, structural proteins and enzymes including human growth hormone, interferon, lysozyme, and β -casein may be produced by the present methods. In one embodiment, proteins are expressed in the oviduct and deposited in eggs of avians, such as chickens, according to the invention. The present invention includes these eggs and these proteins.

The present invention also includes methods of producing transgenic
15 vertebrate animals, for example, transgenic chickens, which employ the use of integrase, cationic polymers and/ nuclear localization signals. The present invention also includes the transgenic vertebrate animals, such as the avians, produced by these methods and other methods disclosed herein. The invention also
20 includes the eggs produced by the transgenic avians produced by these methods and other methods disclosed herein.

In one embodiment, the methods of the invention include introducing into a cell: 1) a nucleic acid comprising a transgene; 2) an integrase activity; and 3) a cationic polymer. Such methods provide for an increased efficiency of transgenic
25 avian production relative to identical methods without the cationic polymer.

In another embodiment, the methods include introducing into a cell: 1) a nucleic acid comprising a transgene; 2) an integrase activity; and 3) a nuclear localization signal. Such methods provide for an increased efficiency of transgenic animal, for example, avian, production relative to identical methods without the
30 nuclear localization signal.

In another embodiment, the methods include introducing into a cell: 1) a nucleic acid comprising a transgene; 2) an integrase activity; 3) a cationic polymer;

and 4) a nuclear localization signal. Such methods provide for an increased efficiency of transgenic vertebrate animal production relative to identical methods without the cationic polymer or without the nuclear localization signal.

In one embodiment, the cell is a cell of an embryo, for example, an avian embryo. In one embodiment, the cell is a cell of an early stage avian embryo comprising a germinal disc. The avian cell may be, for example, a cell of a stage I avian embryo, a cell of a stage II avian embryo, a cell of a stage III avian embryo, a cell of a stage IV avian embryo, a cell of a stage V avian embryo, a cell of a stage VI avian embryo, a cell of a stage VII avian embryo, a cell of a stage VIII avian embryo, a cell of a stage IX avian embryo, a cell of a stage X avian embryo, a cell of a stage XI avian embryo or a cell of a stage XII avian embryo. In one particularly useful embodiment, the avian cell is a cell of a stage X avian embryo. In another useful embodiment, the avian cell is a cell of a stage I avian embryo.

The methods provide for the introduction of nucleic acid into the avian cell by any suitable technique known to those of skill in the art. For example, the nucleic acid may be introduced into the avian cell by microinjecting, transfection, electroporation or lipofection. In one particularly useful embodiment, the introduction of the nucleic acid is accomplished by microinjecting.

The nucleic acid which includes a transgene may be DNA or RNA or a combination of RNA and DNA. The nucleic acid may comprise a single strand or may comprise a double strand. The nucleic acid may be a linear nucleic acid or may be an open or closed circular nucleic acid and may be naturally occurring or synthetic.

Integrase activity may be introduced into the cell, such as an avian cell, in any suitable form. In one embodiment, an integrase protein is introduced into the cell. In another embodiment, a nucleic acid encoding an integrase is introduced into the cell. The nucleic acid encoding the integrase may be double stranded DNA, single stranded DNA, double stranded RNA, single stranded RNA or a single or double stranded nucleic acid which includes both RNA and DNA. In one particularly useful embodiment, the nucleic acid is mRNA. Integrase activity may be introduced into the cell by any suitable technique. Suitable techniques include those described herein for introducing the nucleic acid encoding a transgene into a

cell. In one useful embodiment, the integrase activity is introduced into the cell with the nucleic acid encoding the transgene. For example, the integrase activity may be introduced into the cell in a mixture with the nucleic acid encoding the transgene.

5 In one embodiment, a nuclear localization signal (NLS) is associated with the nucleic acid which includes a transgene. For example, the NLS may be associated with the nucleic acid by a chemical bond. Examples of chemical bonds by which an NLS may be associated with the nucleic acid include an ionic bond, a covalent bond, hydrogen bond and Van der Waal's force. In one particularly
10 useful embodiment, the nucleic acid which includes a transgene is associated with an NLS by an ionic bond. NLS may be introduced into the cell by any suitable technique. Suitable techniques included those described herein for introducing the nucleic acid encoding a transgene into a cell. In one useful embodiment, the NLS is introduced into the cell with the nucleic acid encoding the transgene. For
15 example, the NLS may be introduced into the cell while associated with the nucleic acid encoding the transgene.

 Cationic polymers may be employed to facilitate the production of transgenic vertebrate animals such as avians. For example, the cationic polymers may be employed in combination with integrase and/or NLS. Any suitable cationic
20 polymer may be used. For example, and without limitation, one or more of polyethylenimine, polylysine, DEAE-dextran, starburst dendrimers and starburst polyamidoamine dendrimers may be used. In a particularly useful embodiment, the cationic polymer includes polyethylenimine. The cationic polymer may be introduced into the cell by any suitable technique. Suitable techniques included
25 those described herein for introducing the nucleic acid encoding a transgene into a cell. In one useful embodiment, the cationic polymer is introduced into the cell in a mixture with the nucleic acid encoding the transgene. For example, the cationic polymer may be introduced into the avian cell while associated with the nucleic acid encoding the transgene.

30 In one particularly useful embodiment of the invention, the transgene includes a coding sequence which is expressed in a cell of the transgenic vertebrate animal, for example, a transgenic avian, producing a peptide or a polypeptide (e.g.,

a protein). The coding sequence may be expressed in any or all of the cells of the transgenic animal. For example, the coding sequence may be expressed in the blood, the magnum and/or the sperm of the animal. In a particularly useful embodiment of the invention, the polypeptide is present in an egg, for example, in
5 the egg white, produce by a transgenic avian.

The present invention also includes methods of dispersing nucleic acid in a cell, for example an avian cell (e.g., an avian embryo cell). These methods include introducing into a cell a nucleic acid and a dispersing agent, for example, a cationic polymer (e.g., polyethylenimine, polylysine, DEAE-dextran, starburst dendrimers
10 and/or starburst polyamidoamine dendrimers) in an amount that will disperse the nucleic acid in a cell. Typically, the dispersing of the nucleic acid is a homogeneous dispersing. In one embodiment, the nucleic acid includes a transgene. NLS or integrase activity may also be introduced into the cell.

The methods of the invention include introducing the cell into a recipient
15 animal, for example, an avian such as a chicken, wherein the recipient avian produces an offspring which includes the transgene. The cell may be introduced into a recipient animal by any suitable technique.

The present invention also includes the identification of certain regions in the genome which are advantageous for heterologous gene expression. These
20 regions can be identified by analysis, using methods known in the art, of the transgenic vertebrate animals or cells produced as disclosed herein.

The production of vertebrate animals which are the mature animals developed from the recombinant embryos, ovum and/or sperm of the invention typically are referred to as the G0 generation and are usually hemizygous for each
25 inserted transgene. The G0 generation may be bred to non-transgenic animals to give rise to G1 transgenic offspring which are also hemizygous for the transgene. The G1 hemizygous offspring may be bred to non-transgenic animals giving rise to G2 hemizygous offspring or may be bred together to give rise to G2 offspring homozygous for the transgene. In one embodiment, hemizygotic G2 offspring
30 from the same line can be bred to produce G3 offspring homozygous for the transgene. In one embodiment, hemizygous G0 animals are bred together to give rise to homozygous G1 offspring. These are merely examples of certain useful

breeding schemes. The present invention contemplates the employment of any useful breeding scheme such as those known to individuals of ordinary skill in the art.

Any useful combination of features described herein is included within the scope of the present invention provided that the features included in any such combination are not mutually inconsistent as will be apparent from the context, this specification, and the knowledge of one of ordinary skill in the art. For example, the term transgenic can encompass the term transchromosomal and methodologies useful for transgenic animals (e.g., avians) and cells disclosed herein may also be employed for transchromosomal avians and avian cells.

Additional objects and aspects of the present invention will become more apparent upon review of the detailed description set forth below when taken in conjunction with the accompanying figures, which are briefly described as follows.

Brief Description of the Figures

Fig. 1 illustrates phage integrase-mediated integration. A plasmid vector bearing the transgene includes the attB recognition sequence for the phage integrase. The vector along with integrase-coding mRNA, a vector expressing the integrase, or the integrase protein itself, are delivered into cells or embryos. The integrase recognizes DNA sequences in the avian genome similar to attP sites, termed pseudo-attP, and mediates recombination between the attB and pseudo-attP sites, resulting in the permanent integration of the transgene into the avian genome.

Fig. 2 illustrates the persistent expression of luciferase from a nucleic acid molecule after phiC31 integrase-mediated integration into chicken cells.

Fig. 3 illustrates the results of a puromycin resistance assay to measure phiC31 integrase-mediated integration into chicken cells.

Fig. 4 illustrates phiC31 integrase-mediated integration into quail cells. Puromycin resistance vectors bearing attB sites were cotransfected with phiC31 integrase, or a control vector, into QT6 cells, a quail fibrosarcoma cell line. One day after transfection, puromycin was added. Puromycin resistant colonies were counted 12 days post-transfection.

Figs. 5A and 5B illustrate that phiC31 integrase can facilitate multiple integrations per avian cell. A puromycin resistance vector bearing an attB site was cotransfected with an enhanced green fluorescent protein (EGFP) expression vector bearing an attB site, and a phiC31 integrase expression vector. After puromycin selection, many puromycin resistant colonies expressed EGFP in all of their cells. Figs. 5A and 5B are the same field of view with EGFP illuminated with ultraviolet light (Fig. 5A) and puromycin resistant colonies photographed in visible light (Fig. 5B). In Fig. 5B, there are 4 puromycin resistant colonies, two of which are juxtaposed at the top. One of these colonies expressed EGFP.

Fig. 6 shows maps of the small vectors used for integrase assays.

Fig. 7 shows integrase promotes efficient integration of large transgenes in avian cells.

Fig. 8 shows maps of large vectors used for integrase assays.

Fig. 9 illustrates the nucleotide sequence of the integrase-expressing plasmid pCMV-31int (SEQ ID NO: 1).

Fig. 10 illustrates the nucleotide sequence of the plasmid pCMV-luc-attB (SEQ ID NO: 2).

Fig. 11 illustrates the nucleotide sequence of the plasmid pCMV-luc-attP (SEQ ID NO: 3).

Fig. 12 illustrates the nucleotide sequence of the plasmid pCMV-pur-attB (SEQ ID NO: 4).

Fig. 13 illustrates the nucleotide sequence of the plasmid pCMV-pur-attP (SEQ ID NO: 5).

Fig. 14 illustrates the nucleotide sequence of the plasmid pCMV-EGFP-attB (SEQ ID NO: 6).

Fig. 15 illustrates the nucleotide sequence of the plasmid p12.0-lys-LSPIPMM-CMV-pur-attB (SEQ ID NO: 7).

Fig. 16 illustrates the nucleotide sequence of the plasmid pOMIFN-Ins-CMV-pur-attB (SEQ ID NO: 8).

Fig. 17 illustrates the nucleotide sequence of the integrase-expressing plasmid pRSV-Int (SEQ ID NO: 9).

Fig. 18 illustrates the nucleotide sequence of the plasmid pCR-XL-TOPO-CMV-pur-attB (SEQ ID NO: 10).

Fig. 19 illustrates the nucleotide sequence of the attP containing polynucleotide SEQ ID NO: 11.

5 Fig. 20 illustrates in schematic from the integration of a heterologous att recombination site into an isolated chromosome. The attB sequence is linked to selectable marker such as a puromycin expression cassette and is flanked by sequences found in the target site of the chromosome to be modified. The DNA is transfected into cells containing the chromosome and stable transfectants are
10 selected for by drug resistance. Site specific integration may be confirmed by several techniques including PCR.

Fig. 21 illustrates the persistent expression of luciferase from a nucleic acid molecule after phiC31 integrase-mediated integration into chicken cells bearing a wild-type attP sequence.

15 Fig. 22 illustrates the distribution of plasmid DNA in a stage I embryo.

Fig. 23 illustrates the distribution of plasmid DNA in a stage I embryo in the presence of low molecular weight polyethylenimine.

Fig. 24 illustrates the distribution of plasmid DNA in a stage I embryo in the presence of low molecular weight polyethylenimine.

20 Fig. 25 illustrates the integration of a gene of interest (i.e., OMC24-IRES-EPO) into an artificial chromosome by integration (which takes place inside of a host cell) wherein cells containing the recombinant chromosome can be selected for based on hygromycin resistance.

25 **Definitions and Abbreviations**

For convenience, definitions of certain terms and certain abbreviations employed in the specification, examples and appended claims are collected here.

Abbreviations used in the present specification include the following: aa, amino acid(s); bp, base pair(s); kb, kilobase(s); mb, megabase(s); att, bacterial
30 recombination attachment site; IU, infectious units.

As used in this specification and the appended claims, the singular forms "a," "an" and "the" include plural references unless the content clearly dictates

otherwise. Thus, for example, reference to "an antigen" includes a mixture of two or more such agents.

The term "antibody" as used herein refers to polyclonal and monoclonal antibodies and fragments thereof, and immunologic binding equivalents thereof.

5 Antibodies may include, but are not limited to polyclonal antibodies, monoclonal antibodies (mAbs), humanized or chimeric antibodies, single chain antibodies, Fab fragments, F(ab')₂ fragments, fragments produced by a Fab expression library, anti-idiotypic (anti-Id) antibodies, and epitope-binding fragments of any of the above.

As used herein, an "artificial chromosome" is a nucleic acid molecule that
10 can stably replicate and segregate alongside endogenous chromosomes in a cell. Artificial chromosomes have the capacity to act as gene delivery vehicles by accommodating and expressing foreign genes contained therein. A mammalian artificial chromosome (MAC) refers to chromosomes that have an active mammalian centromere(s). Plant artificial chromosomes, insect artificial
15 chromosomes and avian artificial chromosomes refer to chromosomes that include plant, insect and avian centromeres, respectively. A human artificial chromosome (HAC,) refers to chromosomes that include human centromeres. For exemplary artificial chromosomes, see, for example, U.S. Pat. Nos. 6,025,155, issued February 15, 2000; 6,077,697, issued June 6, 2000; 5,288,625, issued February 22,
20 1994; 5,712,134, issued January 27, 1998; 5,695,967, issued December 9, 1997; 5,869,294, issued February 9, 1999; 5,891,691, issued April 6, 1999 and 5,721,118, issued February 24, 1998 and published International PCT application Nos, WO 97/40183, published October 30, 1997; and WO 98/08964, published March 5, 1998, the disclosures of which are incorporated herein in their entireties by
25 reference. The term "chromosome" may be used interchangeably with the term "artificial chromosome" as will be apparent based on the context of such use.

Foreign genes that can be contained in artificial chromosome expression systems can include, but are not limited to, nucleic acid that encodes therapeutically effective substances, such as anti-cancer agents, enzymes,
30 hormones and antibodies. Other examples of heterologous DNA include, but are not limited to, DNA that encodes traceable marker proteins (reporter genes), such as fluorescent proteins, such as green, blue or red fluorescent proteins (GFP, BFP

and RFP, respectively), other reporter genes, such as beta-galactosidase and proteins that confer drug resistance, such as a gene encoding hygromycin-resistance.

The term "avian" as used herein refers to any species, subspecies or race of organism of the taxonomic class avia, such as, but not limited to chicken, turkey, duck, goose, quail, pheasants, parrots, finches, hawks, crows and ratites including ostrich, emu and cassowary. The term includes the various known strains of Gallus gallus, or chickens, (for example, White Leghorn, Brown Leghorn, Barred-Rock, Sussex, New Hampshire, Rhode Island, Australorp, Minorca, Amrox, California Gray), as well as strains of turkeys, pheasants, quails, duck, ostriches and other poultry commonly bred in commercial quantities. It also includes an individual avian organism in all stages of development, including embryonic and fetal stages. The term "avian" also may denote "pertaining to a bird", such as "an avian (bird) cell."

The terms "chimeric animal" or "mosaic animal" are used herein to refer to an animal in which a nucleotide sequence of interest is found in some but not all cells of the animal, or in which the recombinant nucleic acid is expressed, in some but not all cells of the animal. The term "tissue-specific chimeric animal" indicates that the recombinant gene is present and/or expressed in some tissues but not others.

The term "coding region" as used herein refers to a continuous linear arrangement of nucleotides which may be translated into a polypeptide. A full length coding region is translated into a full length protein; that is, a complete protein as would be translated in its natural state absent any post-translational modifications. A full length coding region may also include any leader protein sequence or any other region of the protein that may be excised naturally from the translated protein.

The term "cytokine" as used herein refers to any secreted polypeptide that affects a function of cells and modulates an interaction between cells in the immune, inflammatory or hematopoietic response. A cytokine includes, but is not limited to, monokines and lymphokines. Examples of cytokines include, but are not limited to, interferon $\alpha 2b$, Interleukin-1 (IL-1), Interleukin-6 (IL-6),

Interleukin-8 (IL-8), Tumor Necrosis Factor- α (TNF- α) and Tumor Necrosis Factor α (TNF- α).

As used herein, "delivery," which is used interchangeably with "transfection," refers to the process by which exogenous nucleic acid molecules are
5 transferred into a cell such that they are located inside the cell.

As used herein, "DNA" is meant to include all types and sizes of DNA molecules including cDNA, plasmids and DNA including modified nucleotides and nucleotide analogs.

The term "expressed" or "expression" as used herein refers to the
10 transcription from a gene to give an RNA nucleic acid molecule at least complementary in part to a region of one of the two nucleic acid strands of the gene. The term "expressed" or "expression" as used herein may also refer to the translation from an RNA molecule to give a protein, a polypeptide or a portion thereof. In one embodiment, for heterologous nucleic acid to be expressed in a
15 host cell, it must initially be delivered into the cell and then, once in the cell, ultimately reside in the nucleus.

The term "gene" or "genes" as used herein refers to nucleic acid sequences that encode genetic information for the synthesis of a whole RNA, a whole protein, or any portion of such whole RNA or whole protein. Genes that are not naturally
20 part of a particular organism's genome are referred to as "foreign genes," "heterologous genes" or "exogenous genes" and genes that are naturally a part of a particular organism's genome are referred to as "endogenous genes". The term "gene product" refers to an RNA or protein that is encoded by the gene. "Endogenous gene products" are RNAs or proteins encoded by endogenous genes.
25 "Heterologous gene products" are RNAs or proteins encoded by "foreign, heterologous or exogenous genes" and are, therefore, not naturally expressed in the cell.

As used herein, the terms "heterologous" and "foreign" with reference to nucleic acids, such as DNA and RNA, are used interchangeably and refer to nucleic
30 acid that does not occur naturally as part of a chromosome, a genome or cell in which it is present or which is found in a location(s) and/or in amounts that differ from the location(s) and/or amounts in which it occurs in nature. It can be nucleic

acid that is not endogenous to the genome, chromosome or cell and has been exogenously introduced into the genome, chromosome or cell. Examples of heterologous DNA include, but are not limited to, DNA that encodes a gene product or gene product(s) of interest, for example, for production of an encoded protein. Examples of heterologous DNA include, but are not limited to, DNA that encodes traceable marker proteins, DNA that encodes therapeutically effective substances, such as anti-cancer agents, enzymes and hormones and as antibodies.

The term "immunoglobulin polypeptide" as used herein refers to a constituent polypeptide of an antibody or a polypeptide derived therefrom. An "immunological polypeptide" may be, but is not limited to, an immunological heavy or light chain and may include a variable region, a diversity region, joining region and a constant region or any combination, variant or truncated form thereof. The term "immunological polypeptides" further includes single-chain antibodies comprised of, but not limited to, an immunoglobulin heavy chain variable region, an immunoglobulin light chain variable region and optionally a peptide linker.

The terms "integrase" and "integrase activity" as used herein refer to a nucleic acid recombinase of the serine recombinase family of proteins.

The term "internal ribosome entry sites (IRES)" as used herein refers to a region of a nucleic acid, most typically an RNA molecule, wherein eukaryotic initiation of protein synthesis occurs far downstream of the 5' end of the RNA molecule. A 43S pre-initiation complex comprising the elf2 protein bound to GTP and Met-tRNA_i^{Met}, the 40S ribosomal subunit, and factors elf3 and 31f1A may bind to an "IRES" before locating an AUG start codon. An "IRES" may be used to initiate translation of a second coding region downstream of a first coding region, wherein each coding region is expressed individually, but under the initial control of a single upstream promoter. An "IRES" may be located in a eukaryotic cellular mRNA.

As used herein, the term "large nucleic acid molecules" or "large nucleic acids" refers to a nucleic acid molecule of at least about 0.05 mb in size, greater than 0.5 mb, including nucleic acid molecules at least about 0.6, 0.7, 0.8, 0.9, 1, 5, 10, 30, 50 and 100, 200, 300, 500 mb in size. Large nucleic acid molecules typically can be on the order of about 10 to about 450 or more mb, and can be of

various sizes, such as, for example, from about 250 to about 400 mb, about 150 to about 200 mb, about 90 to about 120 mb, about 60 to about 100 mb and about 15 to 50 mb. A large nucleic acid molecule may be larger than about 8 kb (e.g., about 8 kb to about 1 mb) as will be apparent based on the context.

5 Examples of large nucleic acid molecules include, but are not limited to, natural chromosomes and fragments thereof, especially mammalian chromosomes and fragments thereof which retain a centromere or retain a centromere and telomeres, artificial chromosome expression systems (ACEs which include a mouse centromere; also called satellite DNA-based artificial chromosomes
10 (SATACs); see U.S. Pat. Nos. 6,025,155, issued February 15; and 6,077,697, issued June 20, 2000), mammalian artificial chromosomes (MACs), plant artificial chromosomes, insect artificial chromosomes, avian artificial chromosomes and minichromosomes (see, e.g., U.S. Pat. Nos. 5,712,134, issued January 27, 1998; 5,891,691, issued April 6, 1999; and 5,288,625, issued February 22, 1994). Useful
15 large nucleic acid molecules can include a single copy of a desired nucleic acid fragment encoding a particular nucleotide sequence, such as a gene of interest, or can carry multiple copies thereof or multiple genes or different heterologous sequences of nucleotides. For example, the chromosomes may carry 1 to about 100 or 1 to about 1000 or even more copies of a gene of interest. Large nucleic acid
20 molecules can be associated with proteins, for example chromosomal proteins, that typically function to regulate gene expression and/or participate in determining overall structure.

 A "nucleic acid fragment of interest" or "nucleotide sequence of interest" may be a trait-producing sequence, by which it is meant a sequence conferring a
25 non-native trait upon the cell in which the protein encoded by the trait-producing sequence is expressed. The term "non-native" when used in the context of a trait-producing sequence means that the trait produced is different than one would find in an unmodified organism which can mean that the organism produces high amounts of a natural substance in comparison to an unmodified organism, or
30 produces a non-natural substance. For example, the genome of a bird could be modified to produce proteins not normally produced in birds such as, for example,

useful animal proteins (e.g., human proteins) such as hormones, cytokines and antibodies.

A nucleic acid fragment of interest may additionally be a "marker nucleic acid" or expressed as a "marker polypeptide". Marker genes encode proteins that can be easily detected in transformed cells and are, therefore, useful in the study of those cells. Examples of suitable marker genes include β -galactosidase, green or yellow fluorescent proteins, enhanced green fluorescent protein, chloramphenicol acetyl transferase, luciferase, and the like. Such regions may also include those 5' noncoding sequences involved with initiation of transcription and translation, such as the enhancer, TATA box, capping sequence, CAAT sequence, and the like.

As used herein, "nucleic acid" refers to a polynucleotide containing at least two covalently linked nucleotide or nucleotide analog subunits. A nucleic acid can be a deoxyribonucleic acid (DNA), a ribonucleic acid (RNA), or an analog of DNA or RNA. Nucleotide analogs are commercially available and methods of preparing polynucleotides containing such nucleotide analogs are known (Lin et al. (1994) Nucl. Acids Res. 22:5220-5234; Jellinek et al. (1995) Biochemistry 34:11363-11372; Pagratis et al. (1997) Nature Biotechnol. 15:68-73). The nucleic acid can be single-stranded, double-stranded, or a mixture thereof. For purposes herein, unless specified otherwise, the nucleic acid is double-stranded, or if it is apparent from the context that the nucleic acid is not double stranded. Nucleic acids include any natural or synthetic linear and sequential array of nucleotides and nucleosides, for example cDNA, genomic DNA, mRNA, tRNA, oligonucleotides, oligonucleosides and derivatives thereof. For ease of discussion, certain nucleic acids may be collectively referred to herein as "constructs," "plasmids," or "vectors."

Techniques useful for isolating and characterizing the nucleic acids and proteins of the present invention are well known to those of skill in the art and standard molecular biology and biochemical manuals may be consulted to select suitable protocols without undue experimentation. See, for example, Sambrook et al, 1989, "Molecular Cloning: A Laboratory Manual", 2nd ed., Cold Spring Harbor, the content of which is herein incorporated by reference in its entirety.

A "nucleoside" is conventionally understood by workers of skill in fields related to the present invention as comprising a monosaccharide linked in glycosidic linkage to a purine or pyrimidine base. A "nucleotide" comprises a nucleoside with at least one phosphate group appended, typically at a 3' or a 5' position (for pentoses) of the saccharide, but may be at other positions of the saccharide. A nucleotide may be abbreviated herein as "nt." Nucleotide residues occupy sequential positions in an oligonucleotide or a polynucleotide. Accordingly a modification or derivative of a nucleotide may occur at any sequential position in an oligonucleotide or a polynucleotide. All modified or derivatized oligonucleotides and polynucleotides are encompassed within the invention and fall within the scope of the claims. Modifications or derivatives can occur in the phosphate group, the monosaccharide or the base.

By way of nonlimiting examples, the following descriptions provide certain modified or derivatized nucleotides. The phosphate group may be modified to a thiophosphate or a phosphonate. The phosphate may also be derivatized to include an additional esterified group to form a triester. The monosaccharide may be modified by being, for example, a pentose or a hexose other than a ribose or a deoxyribose. The monosaccharide may also be modified by substituting hydroxyl groups with hydro or amino groups, by esterifying additional hydroxyl groups. The base may be modified as well. Several modified bases occur naturally in various nucleic acids and other modifications may mimic or resemble such naturally occurring modified bases. Nonlimiting examples of modified or derivatized bases include 5-fluorouracil, 5-bromouracil, 5-chlorouracil, 5-iodouracil, hypoxanthine, xanthine, 4-acetylcytosine, 5-(carboxyhydroxymethyl) uracil, 5-carboxymethylaminomethyl-2-thiouridine, 5-carboxymethylaminomethyluracil, dihydrouracil, beta-D-galactosylqueosine, inosine, N6-isopentenyladenine, 1-methylguanine, 1-methylinosine, 2,2-dimethylguanine, 2-methyladenine, 2-methylguanine, 3-methylcytosine, 5-methylcytosine, N6-adenine, 7-methylguanine, 5-methylaminomethyluracil, 5-methoxyaminomethyl-2-thiouracil, beta-D-mannosylqueosine, 5'-methoxycarboxymethyluracil, 5-methoxyuracil, 2-methylthio-N6-isopentenyladenine, uracil-5-oxyacetic acid (v), wybutoxosine, pseudouracil, queosine, 2-thiocytosine, 5-methyl-2-thiouracil, 2-thiouracil, 4-

thiouracil, 5-methyluracil, uracil-5-oxyacetic acid methylester, uracil-5-oxyacetic acid (v), 5-methyl-2-thiouracil, 3-(3-amino-3-N-2-carboxypropyl) uracil, (acp3)w, and 2,6-diaminopurine. Nucleotides may also be modified to harbor a label. Nucleotides may also bear a fluorescent label or a biotin label.

5 The term “operably linked” refers to an arrangement of elements wherein the components so described are configured so as to perform their usual function. Control sequences operably linked to a coding sequence are capable of effecting the expression of the coding sequence. The control sequences need not be contiguous with the coding sequence, so long as they function to direct the
10 expression thereof. For example, intervening untranslated yet transcribed sequences can be present between a promoter sequence and the coding sequence and the promoter sequence can still be considered “operably linked” to the coding sequence.

 “Therapeutic proteins” or “pharmaceutical proteins” include an amino acid
15 sequence which in whole or in part makes up a drug. In one embodiment, a pharmaceutical composition or therapeutic composition includes one or more pharmaceutical proteins or therapeutic proteins.

 The terms “polynucleotide,” “oligonucleotide,” and “nucleic acid sequence” are used interchangeably herein and include, but are not limited to, coding
20 sequences (polynucleotide(s) or nucleic acid sequence(s) which are transcribed and translated into polypeptide in vitro or in vivo when placed under the control of appropriate regulatory or control sequences); control sequences (e.g., translational start and stop codons, promoter sequences, ribosome binding sites, polyadenylation signals, transcription factor binding sites, transcription termination sequences,
25 upstream and downstream regulatory domains, enhancers, silencers, and the like); and regulatory sequences (DNA sequences to which a transcription factor(s) binds and alters the activity of a gene’s promoter either positively (induction) or negatively (repression)). No limitation as to length or to synthetic origin are suggested by the terms described above.

30 As used herein the terms “peptide,” “polypeptide” and “protein” refer to a polymer of amino acids in a serial array, linked through peptide bonds. A “peptide” typically is a polymer of at least two to about 30 amino acids linked in a

serial array by peptide bonds. The term "polypeptide" includes proteins, protein fragments, protein analogues, oligopeptides and the like. The term "polypeptides" contemplates polypeptides as defined above that are encoded by nucleic acids, produced through recombinant technology (isolated from an appropriate source
5 such as a bird), or synthesized. The term "polypeptides" further contemplates polypeptides as defined above that include chemically modified amino acids or amino acids covalently or noncovalently linked to labeling moieties.

The terms "percent sequence identity" or "percent sequence similarity" as used herein refer to the degree of sequence identity between two nucleic acid
10 sequences or two amino acid sequences as determined using the algorithm of Karlin & Attschul, Proc. Natl. Acad. Sci. 87: 2264-2268 (1990), modified as in Karlin & Attschul, Proc. Natl. Acad. Sci. 90: 5873-5877 (1993). Such an algorithm is incorporated into the NBLAST and XBLAST programs of Attschul et al, 1990, T. Mol. Biol. 215: 403-410. BLAST nucleotide searches are performed
15 with the NBLAST program, score = 100, word length = 12, to obtain nucleotide sequences homologous to a nucleic acid molecule of the invention. BLAST protein searches are performed with the XBLAST program, score = 50, word length = 3, to obtain amino acid sequences homologous to a reference polypeptide. To obtain gapped alignments for comparison purposes, Gapped BLAST is utilized
20 as described in Attschul et al, Nucl. Acids Res. 25: 3389-3402 (1997). When utilizing BLAST and Gapped BLAST programs, the default parameters of the respective programs (e.g. XBLAST and NBLAST) are used. Other algorithms, programs and default settings may also be suitable such as, but not only, the GCG-Sequence Analysis Package of the U.K. Human Genome Mapping Project
25 Resource Centre that includes programs for nucleotide or amino acid sequence comparisons. Examples of useful algorithms are FASTA and BESTFIT.

The term "promoter" as used herein refers to the DNA sequence that determines the site of transcription initiation by an RNA polymerase. A "promoter-proximal element" is a regulatory sequence generally within about 200
30 base pairs of the transcription start site.

The term "pseudo-recombination site" as used herein refers to a site at which an integrase can facilitate recombination even though the site may not have a

sequence identical to the sequence of its wild-type recombination site. For example, a phiC31 integrase and vector carrying a phiC31 wild-type recombination site can be placed into an avian cell. The wild-type recombination sequence aligns itself with a sequence in the avian cell genome and the integrase facilitates a recombination event. When the sequence from the genomic site in the avian cell, where the integration of the vector took place, is examined, the sequence at the genomic site typically has some identity to, but may not be identical with, the wild-type bacterial genome recombination site. The recombination site in the avian cell genome is considered to be a pseudo-recombination site (e.g., a pseudo-attP site) at least because the avian cell is heterologous to the normal phiC31 phage/bacterial cell system. The size of the pseudo-recombination site can be determined through the use of a variety of methods including, but not limited to, (i) sequence alignment comparisons, (ii) secondary structural comparisons, (iii) deletion or point mutation analysis to find the functional limits of the pseudo-recombination site, and (iv) combinations of the foregoing.

The terms "recombinant cell" and "genetically transformed cell" refer to a cell comprising a combination of nucleic acid segments not found in a single cell with each other in nature. A new combination of nucleic acid segments can be introduced into an organism using a wide array of nucleic acid manipulation techniques available to those skilled in the art. The recombinant cell may harbor a vector that is extragenomic, i.e. that does not covalently insert into the cellular genome, including a non-nuclear (e.g. mitochondrial) genome(s). A recombinant cell may further harbor a vector or a portion thereof that is intragenomic, i.e. covalently incorporated within the genome of the recombinant cell.

The term "recombination site" as used herein refers to a polynucleotide stretch comprising a recombination site normally recognized and used by an integrase. For example, λ phage is a temperate bacteriophage that infects *E. coli*. The phage has one attachment site for recombination (attP) and the *E. coli* bacterial genome has an attachment site for recombination (attB). Both of these sites are recombination sites for λ integrase. Recombination sites recognized by a particular integrase can be derived from a homologous system and associated with

heterologous sequences, for example, the attP site can be placed in other systems to act as a substrate for the integrase.

The terms "recombinant nucleic acid" and "recombinant DNA" as used herein refer to combinations of at least two nucleic acid sequences that are not
5 naturally found in a eukaryotic or prokaryotic cell. The nucleic acid sequences may include, but are not limited to, nucleic acid vectors, gene expression regulatory elements, origins of replication, suitable gene sequences that when expressed confer antibiotic resistance, protein-encoding sequences and the like. The term "recombinant polypeptide" is meant to include a polypeptide produced by
10 recombinant DNA techniques. A recombinant polypeptide may be distinct from a naturally occurring polypeptide either in its location, purity or structure. Generally, a recombinant polypeptide will be present in a cell in an amount different from that normally observed in nature.

As used herein, the term "satellite DNA-based artificial chromosome
15 (SATAC)" (e.g., ACE) is a type of artificial chromosome. These artificial chromosomes are substantially all neutral non-coding sequences (heterochromatin) except for foreign heterologous, typically gene-encoding nucleic acid, that is present within (see U.S. Pat. Nos. 6,025,155, issued February 15, 2000 and 6,077,697, issued June 20, 2000 and International PCT application No. WO
20 97/40183, published October 30, 1997).

The term "source of integrase activity" as used herein refers to a polypeptide or multimeric protein having serine recombinase (integrase) activity in an avian cell. The term may further refer to a polynucleotide encoding the serine recombinase, such as an mRNA, an expression vector, a gene or isolated gene that
25 may be expressed as the recombinase-specific polypeptide or protein.

As used herein the term "therapeutic substance" refers to a component that comprises a substance which can provide for a therapeutic effect, for example, a therapeutic protein.

"Transchromosomic avian" means an avian which contains an artificial
30 chromosome in some or all of its cells. A transchromosomic avian can include the artificial chromosome in its germ cells.

The term "transcription regulatory sequences" as used herein refers to nucleotide sequences that are associated with a gene nucleic acid sequence and which regulate the transcriptional expression of the gene. Exemplary transcription regulatory sequences include enhancer elements, hormone response elements, steroid response elements, negative regulatory elements, and the like.

The term "transfection" as used herein refers to the process of inserting a nucleic acid into a host cell. Many techniques are well known to those skilled in the art to facilitate transfection of a nucleic acid into an eukaryotic cell. These methods include, for instance, treating the cells with high concentrations of salt such as a calcium or magnesium salt, an electric field, detergent, or liposome mediated transfection, to render the host cell competent for the uptake of the nucleic acid molecules, and by such methods as micro-injection into a pro-nucleus, sperm-mediated and restriction-mediated integration.

The term "transformed" as used herein refers to a heritable alteration in a cell resulting from the uptake of a heterologous DNA.

As used herein, the term "transgene" means a nucleic acid sequence that is partly or entirely heterologous, i.e., foreign, to the transgenic animal or cell into which it is introduced, or, is homologous to an endogenous gene of the transgenic animal or cell into which it is introduced, but which is designed to be inserted, or is inserted, into the animal's genome in such a way as to alter the genome of the cell into which it is inserted (e.g., it is inserted at a location which differs from that of the natural gene or its insertion results in a knockout).

As used herein, a "transgenic avian" is any avian, as defined herein, in which one or more of the cells of the avian contain heterologous nucleic acid introduced by manipulation, such as by transgenic techniques. The nucleic acid may be introduced into a cell, directly or indirectly, by introduction into a precursor of the cell by way of deliberate genetic manipulation, such as by microinjection or by infection with a recombinant virus. Genetic manipulation also includes classical cross-breeding, or in vitro fertilization. A recombinant DNA molecule may be integrated within a chromosome, or it may be extrachromosomally replicating DNA.

The term "trisomic" as used herein refers to a cell or animal, such as an avian cell or bird that has a $2n+1$ chromosomal complement, where n is the haploid number of chromosomes, for the animal species concerned.

5 The terms "vector" or "nucleic acid vector" as used herein refer to a natural or synthetic single or double stranded plasmid or viral nucleic acid molecule (RNA or DNA) that can be transfected or transformed into cells and replicate independently of, or within, the host cell genome. The term "expression vector" as used herein refers to a nucleic acid vector that comprises a transcription regulatory region operably linked to a site wherein is, or can be, inserted, a nucleotide
10 sequence to be transcribed and, optionally, to be expressed, for instance, but not limited to, a sequence coding at least one polypeptide.

Detailed Description

The present invention provides for recombinant vertebrate cells (e.g.,
15 transgenic or transchromosomal avian cells) and transgenic vertebrate animals (e.g., transgenic or transchromosomal avians) and methods of making the cells and the animals. For example, the invention provides for methods of inserting nucleotide sequences into the genome of vertebrate animals or into the cells of vertebrate animals in a site specific manner. Examples of vertebrates include,
20 without limitation, birds, mammals, fish, reptiles and amphibians. Examples of mammals include sheep, goats and cows. In one certain embodiment of the invention, the vertebrate animals are birds or avians. Examples of birds include, without limitation, chickens, turkeys, ducks, geese, quail, pheasants, parrots, finches, hawks, crows and ratites including ostriches, emu and cassowary.

25 In one embodiment, the present invention provides for methods of inserting nucleotide sequences into the genome of an animal using methods of transgenesis based on site specific integration, for example, site specific integrase mediated-transgenesis. The present invention contemplates any useful method of integrase mediated transgenesis including but not limited to, transgenesis mediated by serine
30 recombinases and tyrosine recombinases. Serine recombinases are well known in the art and include without limitation, EcoYBCK, Φ C31, SCH10.38c, SCC88.14, SC8F4.15c, SCD12A.23, Bxb1, WwK, Sau CcrB, Bsu CisB, TP901-1, Φ 370.1,

5 Φ 105, Φ FC1, A118, Cac1956, Cac1951, Sau CcrA, Spn, TnpX, TndX, SPBc2, SC3C8.24, SC2E1.37, SCD78.04c, R4, Φ Rv1, Y4bA, Bja, SsoISC1904b, SsoISC1904a, Aam, MjaMJ1004, Pab, SsoISC1913, HpyIS607, MceRv0921, MtuRv0921, MtuRv2979c, MtuRv2792c, MtuISY349, MtuRv3828c, SauSK1,
 10 Spy, EcoTn21, Mlo92, EcoTn3, Lla, Cpe, SauSK41, BmeTn5083, SfaTn917, Bme53, Ran, RmzY4CG, SarpNL1, Pje, Xan, ISXc5, Pae, Xca, Req, Mlo90, PpsTn5501, pMER05, Cgl, MuGin, StyHin, Xfa911, Xfa910, Rrh, SauTn552 and Aac serine recombinases. Tyrosine recombinases well known in the art include without limitation, BS codV, BS ripX, BS ydcL, CB tnpA, Col1D, CP4, Cre, D29,
 15 DLP12, DN int, EC FimB, EC FimE, EC orf, EC xerC, EC xerD, Φ 11, Φ 13, Φ 80, Φ adh, Φ CTX, Φ LC3, FLP, Φ R73, Hlorf, HI rci, HI xerC, HI xerD, HK22, HP1, L2, L5, L54, λ , LL orf, LL xerC, LO L5, MJ orf, ML orf, MP int, MT int, MT orf, MV4, P186, P2, P21, P22, P4, P434, PA sss, PM fimB, pAE1, pCL1, pKD1, pMEA, pSAM2, pSB2, pSB3, pSDL2, pSE101, pSE211, pSM1, pSR1, pWS58,
 20 R721, Rci, SF6, SLP1, SM orf, SsrA, SSV1, T12, Tn21, Tn4430, Tn554a, Tn554b, Tn7, Tn916, Tuc, WZ int, XisA and XisC. Other enzymes which may be useful for mediation of transgenesis in accordance with the present invention include, certain transposases, invertases and resolvases.

25 In certain instances, integration host factors (IHF) may be necessary for the integration of nucleotide sequences of the invention into the genome of cells as disclosed herein. In such a case, the integration host factors may be delivered to the cells directly or they may be delivered to the cells in the form of a nucleic acid which, in the case of RNA, is translated to produce the IHF or, in the case of DNA, is transcribed and translated to produce the IHF.

30 The present invention contemplates the use of any system capable of site specifically inserting a nucleotide sequence of interest into the genome of a cell, for example, to produce a transgenic vertebrate animal. Typically, although not exclusively, these systems require at least three components: 1) a sequence in the genome which specifies the site of insertion; 2) a nucleotide sequence which is directed to the site of insertion and an enzyme which catalyzes the insertion of the nucleotide sequence into the genome at the site of insertion. Many enzymes, including integrases, which are capable of site specifically inserting nucleotide

sequences into the genome have been characterized. Examples of these enzymes are disclosed in for example, Esposito et al (1997) Nucleic Acids Research, 25;3605-3614 and Nunes-Düby et al (1998) Nucleic Acids Research, 26; 391-406. The disclosure of each of these references is incorporated herein in their entirety.

5 In one embodiment of the present invention, a serine recombinase is employed. Serine recombinase integrase mediates recombination between an attB site on a transgene vector and an attP or a pseudo attP site on a chromosome. In the method of the invention for integrase-mediated transgenesis, a heterologous wild-type attP site can be integrated into a nuclear genome to create a transgenic
10 cell line or a transgenic vertebrate animal, such as an avian. A serine recombinase (integrase) and an attB-bearing transgene vector are then introduced into cells harboring the heterologous attP site, or into embryos derived from animals which bear the attP recombination site. The locations of attP and attB may be reversed such that the attB site is inserted into a chromosome and the attP sequence resides
15 in an incoming transgene vector. In either case, the att site of the introduced vector would then preferentially recombine with the integrated heterologous att site in the genome of the recipient cell.

The methods of the invention are based, in part, on the discovery that there exists in vertebrate animal genomes, such as avian genomes, a number of specific
20 nucleic acid sequences, termed pseudo-recombination sites, the sequences of which may be distinct from wild-type recombination sites but which can be recognized by a site-specific integrase and used to promote the efficient insertion of heterologous genes or polynucleotides into the targeted nuclear genome. The inventors have identified pseudo-recombination sites in avian cells capable of recombining with a
25 recombination site, such as an attB site within a recombinant nucleic acid molecule introduced into the target avian cell. The invention is also based on the prior integration of a heterologous att recombination site, typically isolated from a bacteriophage or a modification thereof, into the genome of the target avian cell.

Integration into a predicted chromosomal site is useful to improve the
30 predictability of expression, which is particularly advantageous when creating transgenic avians. Transgenesis by methods that result in insertion of the transgene

into random positions of the avian genome is unpredictable since the transgene may not express at the expected levels or in the predicted tissues.

The invention as disclosed herein, therefore, provides methods for site-specifically genetically transforming an avian nuclear genome. In general, an avian
5 cell having a first recombination site in the nuclear genome is transformed with a site-specific polynucleotide construct comprising a second recombination sequence and one or more polynucleotides of interest. Into the same cell, integrase activity may be introduced that specifically recognizes the first and second recombination sites under conditions such that the polynucleotide sequence of interest is inserted
10 into the nuclear genome via an integrase-mediated recombination event between the first and second recombination sites.

The integrase activity, or a source thereof, can be introduced into the cell prior to, or concurrent with, the introduction of the site-specific construct. The integrase can be delivered to a cell as a polypeptide, or by expressing the integrase
15 from a source polynucleotide such as an mRNA or from an expression vector that encodes the integrase, either of which can be delivered to the target cell before, during or after delivery of the polynucleotide of interest. Any integrase that has activity in a cell may be useful in the present invention, including HK022 (Kolot et al, Biotechnol. Bioeng., 84: 56-60 (2003)). In one embodiment, the integrase is a
20 serine recombinase as described, for example, by Smith & Thorpe, in Mol. Microbiol., 44: 299-307 (2002). For example, the integrase may be TP901-1 (Stoll et al, J. Bact., 184: 3657-3663 (2002); Olivares et al, Gene, 278:167-176 (2001) or the integrase from the phage phiC31.

The nucleotide sequence of the junctions between an integrated transgene
25 into the attP (or attB site) would be known. Thus, a PCR assay can be designed by one of skill in the art to detect when the integration event has occurred. The PCR assay for integration into a heterologous wild-type attB or attP site can also be readily incorporated into a quantitative PCR assay using TAQMAN™ or related technology so that the efficiency of integration can be measured.

30 In one embodiment, the minimal attB and attP sites able to catalyze recombination mediated by the phiC31 integrase are 34 and 39 bp, respectively. In cell lines that harbor a heterologous integrated attP site, however, integrase may

have a preference for the inserted attP over any pseudo-attP sites of similar length, because pseudo-attP sites have very low sequence identity (for example, between 10 to 50% identity) compared to the more efficient wild-type attP sequence. It is within the scope of the methods of the invention, however, for the recombination site within the target genome to be a pseudo-att site such as a pseudo-attP site or an attP introduced into a genome.

The sites used for recognition and recombination of phage and bacterial DNAs (the native host system) are generally non-identical, although they typically have a common core region of nucleic acids. In one embodiment, the bacterial sequence is called the attB sequence (bacterial attachment) and the phage sequence is called the attP sequence (phage attachment). Because they are different sequences, recombination can result in a stretch of nucleic acids (for example, attL or attR for left and right) that is neither an attB sequence or an attP sequence, and likely is functionally unrecognizable as a recombination site to the relevant enzyme, thus removing the possibility that the enzyme will catalyze a second recombination reaction that would reverse the first.

The integrase may recognize a recombination site where sequence of the 5' region of the recombination site can differ from the sequence of the 3' region of the recombination sequence. For example, for the phage phiC31 attP (the phage attachment site), the core region is 5'-TTG-3' the flanking sequences on either side are represented here as attP5' and attP3', the structure of the attP recombination site is, accordingly, attP5'-TTG-attP3'. Correspondingly, for the native bacterial genomic target site (attB) the core region is 5'-TTG-3', and the flanking sequences on either side are represented here as attB5' and attB3', the structure of the attB recombination site is, accordingly, attB5'-TTG-attB3'. After a single-site, phiC31 integrase-mediated recombination event takes place between the phiC31 phage and the bacterial genome, the result is the following recombination product: attB5'-TTG-attP3'{phiC31 vector sequences}-attP5'-TTG-attB3'. In the method of invention, the attB site will be within a recombinant nucleic acid molecule that may be delivered to a target cell. The corresponding attP (or pseudo-attP) site will be within the cell nuclear genome. Consequently, after phiC31 integrase mediated recombination, the recombination product, the nuclear genome with the integrated

heterologous polynucleotide will have the sequence attP5'-TTG-attB3'{heterologous polynucleotide}-attB5'-TTG-attP3'. Typically, after recombination the post-recombination recombination sites are no longer able to act as substrate for the phiC31 integrase. This results in stable integration with little or
5 no integrase mediated excision.

While the one useful recombination site to be included in the recombinant nucleic acid molecules and modified chromosomes of the present invention is the attP site, it is contemplated that any attP-like site may be used if compatible with the attB site. For instance, any pseudo-attP site of the chicken genome may be
10 identified according to the methods of Example 7 herein and used as a heterologous att recombination site. For example, such attP-like sites may have a sequence that is greater than at least 25% identical to SEQ ID NO: 11 as shown in Fig. 19, such as described in Groth et al, Proc. Natl. Acad. Sci. U.S.A. 97: 5995-6000 (2000) incorporated herein by reference in its entirety. In one embodiment,
15 the selected site will have a similar degree of efficiency of recombination, for example, at least the same degree of efficiency of recombination as the attP site (SEQ ID NO: 11) itself.

In the present invention, the recipient cell population may be an isolated cell line such as, for example, DF-1 chicken fibroblasts, chicken DT40 cells or a
20 cell population derived from an early stage embryo, such as a chicken stage I embryo or mid stage or late stage (e.g., stage X) embryos. One useful avian cell population is blastodermal cells isolated from a stage X avian embryo. The methods of the present invention, therefore, include steps for the isolation of blastodermal cells that are then suspended in a cell culture medium or buffer for
25 maintaining the cells in a viable state, and which allows the cell suspension to contact the nucleic acids of the present invention. It is also within the scope of the invention for the nucleic acid construct and the source of integrase activity to be delivered directly to an avian embryo such as a blastodermal layer, or to a tissue layer of an adult bird such as the lining of an oviduct.

30 When the recipient cell population is isolated from an early stage avian embryo, the embryos must first be isolated. For stage I avian embryos from, for example, a chicken, a fertilized ovum is surgically removed from a bird before the

deposition of the outer hard shell has occurred. The nucleic acids for integrating a heterologous nucleic acid into a recipient cell genome may then be delivered to isolated embryos by lipofection, microinjection (as described in Example 6 below) or electroporation and the like. After delivery of the nucleic acid, the transfected embryo and its yolk may be deposited into the infundibulum of a recipient hen for the deposition of egg white proteins and a hard shell, and laying of the egg. Stage X avian embryos are obtained from freshly laid fertilized eggs and the blastodermal cells isolated as a suspension of cells in a medium, as described in Example 4 below. Isolated stage X blastodermal cell populations, once transfected, may be injected into recipient stage X embryos and the hard shell eggs resealed according to the methods described in U.S. Patent No. 6,397,777, issued June 4, 2002, the disclosure of which is incorporated in its entirety by reference herein.

In one embodiment of the invention, once a heterologous nucleic acid is delivered to the recipient cell, integrase activity is expressed. The expressed integrase (or injected integrase polypeptide) then mediates recombination between the att site of the heterologous nucleic acid molecule, and the att (or pseudo att) site within the genomic DNA of the recipient avian cell.

It is within the scope of the present invention for the integrase-encoding sequence and a promoter operably linked thereto to be included in the delivered nucleic acid molecule and that expression of the integrase activity occurs before integration of the heterologous nucleic acid into the cell genome. In one embodiment, an integrase-encoding nucleic acid sequence and associated promoter are in an expression vector that may be co-delivered to the recipient cell with the heterologous nucleic acid molecule to be integrated into the recipient genome.

One suitable integrase expressing expression vector for use in the present invention is pCMV-C31int (SEQ ID NO: 1) as shown in Fig. 9, and described in Groth et al, Proc. Natl. Acad. Sci. U.S.A. 97: 5995-6000 (2000), incorporated herein by reference in its entirety. In pCMV-C31int, expression of the integrase-encoding sequence is driven by the CMV promoter. However, any promoter may be used that will give expression of the integrase in a recipient cell, including operably linked avian-specific gene expression control regions of the avian

ovalbumin, lysozyme, ovomucin, ovomucoid gene loci, viral gene promoters, inducible promoters, the RSV promoter and the like.

The recombinant nucleic acid molecules of the present invention for delivery of a heterologous polynucleotide to the genome of a recipient cell may
5 comprise a nucleotide sequence encoding the attB attachment site of *Streptomyces* ambotfaciens as described in Thorpe & Smith, Proc. Natl. Acad. Sci. U.S.A. 95: 5505-5510 (1998). The nucleic acid molecule of the present invention may further comprise an expression cassette for the expression in a recipient cell of a heterologous nucleic acid encoding a desired heterologous polypeptide.
10 Optionally, the nucleic acid molecules may also comprise a marker such as, but not limited to, a puromycin resistance gene, a luciferase gene, EGFP, and the like.

It is contemplated that the expression cassette, for introducing a desired heterologous polypeptide, comprises a promoter operably linked to a nucleic acid encoding the desired polypeptide and, optionally, a polyadenylation signal
15 sequence. Exemplary nucleic acids suitable for use in the present invention are more fully described in the examples below.

In one embodiment of the present invention, following delivery of the nucleic acid molecule and a source of integrase activity into a cell population, for example, an avian cell population, the cells are maintained under culture conditions
20 suitable for the expression of the integrase and/or for the integrase to mediate recombination between the recombination site of the nucleic acid and recombination site in the genome of a recipient cell. When the recipient cell is cultured in vitro, such cells may be incubated at 37° Celsius. For example, chicken early stage blastodermal cells may be incubated at 37° Celsius. They may then be
25 injected into an embryo within a hard shell, which is resealed for incubation until hatching. Alternatively, the transfected cells may be maintained in in vitro culture.

In one embodiment, the present invention provides methods for the site-specific insertion of a heterologous nucleic acid molecule into the nuclear genome of a cell by delivering to a target cell that has a recombination site in its nuclear
30 genome, a source of integrase activity, a site-specific construct that has another recombination site and a polynucleotide of interest, and allowing the integrase activity to facilitate a recombination event between the two recombination sites,

thereby integrating the polynucleotide of interest into the nuclear genome.

(a) Expression vector nucleic acid molecules: A variety of recombinant nucleic acid expression vectors are suitable for use in the practice of the present invention. The site-specific constructs described herein can be constructed utilizing
5 methodologies well known in the art of molecular biology (see, for example, Ausubel or Maniatis) in view of the teachings of the specification. As described above, the constructs are assembled by inserting into a suitable vector backbone a recombination site such as an attP or an attB site, a polynucleotide of interest operably linked to a gene expression control region of interest and, optionally a
10 sequence encoding a positive selection marker. Polynucleotides of interest can include, but are not limited to, expression cassettes encoding a polypeptide to be expressed in the transformed cell or in a transgenic vertebrate animal derived therefrom. The site-specific constructs are typically, though not exclusively, circular and may also contain selectable markers, an origin of replication, and other
15 elements.

Any of the vectors of the present invention may also optionally include a sequence encoding a signal peptide that directs secretion of the polypeptide expressed by the vector from the transgenic cells, for instance, from tubular gland cells of the oviduct of an avian. In one embodiment, this aspect of the invention
20 effectively broadens the spectrum of exogenous proteins that may be deposited in the whites of avian eggs using the methods of the invention. Where an exogenous polypeptide would not otherwise be secreted, the vector bearing the coding sequence can be modified to comprise, for instance, about 60 bp encoding a signal peptide. The DNA sequence encoding the signal peptide may be inserted in the
25 vector such that the signal peptide is located at the N-terminus of the polypeptide encoded by the vector.

The expression vectors of the present invention can comprise a transcriptional regulatory region, for example, an avian transcriptional regulatory region, for directing expression of either fusion or non-fusion proteins. With
30 fusion vectors, a number of amino acids are usually added to the desired expressed target gene sequence such as, but not limited to, a polypeptide sequence for thioredoxin. A proteolytic cleavage site may further be introduced at a site

between the target recombinant protein and the fusion sequence. Additionally, a region of amino acids such as a polymeric histidine region may be introduced to allow binding of the fusion protein to metallic ions such as nickel bonded to a solid support, for purification of the fusion protein. Once the fusion protein has been
5 purified, the cleavage site allows the target recombinant protein to be separated from the fusion sequence. Enzymes suitable for use in cleaving the proteolytic cleavage site include, but are not limited to, Factor Xa and thrombin. Fusion expression vectors that may be useful in the present invention include pGex (Amrad Corp., Melbourne, Australia), pRIT5 (Pharmacia, Piscataway, NJ) and
10 pMAL (New England Biolabs, Beverly, MA), that fuse glutathione S-transferase, protein A, or maltose E binding protein, respectively, to a desired target recombinant protein.

Epitope tags are short peptide sequences that are recognized by epitope specific antibodies. A fusion protein comprising a recombinant protein and an
15 epitope tag can be simply and easily purified using an antibody bound to a chromatography resin, for example. The presence of the epitope tag furthermore allows the recombinant protein to be detected in subsequent assays, such as Western blots, without having to produce an antibody specific for the recombinant protein itself. Examples of commonly used epitope tags include V5, glutathione-S-
20 transferase (GST), hemagglutinin (HA), the peptide Phe-His-His-Thr-Thr, chitin binding domain, and the like.

Exemplary gene expression control regions for use in cells such as avian cells (e.g., chicken cells) include, but are not limited to, avian specific promoters such as the chicken lysozyme, ovalbumin, or ovomucoid promoters, and the like.
25 Particularly useful in avian systems are tissue-specific promoters such as avian oviduct promoters that allow for expression and delivery of a heterologous polypeptide to an egg white.

Viral promoters serve the same function as bacterial or eukaryotic promoters and either provide a specific RNA polymerase in trans (bacteriophage
30 T7) or recruit cellular factors and RNA polymerase (SV40, RSV, CMV). Viral promoters can be useful as they are generally particularly strong promoters. One useful promoter for employment in avian cells is the RSV promoter.

Selection markers are valuable elements in expression vectors as they provide a means to select for growth of only those cells that contain a vector. Common selectable marker genes include those for resistance to antibiotics such as ampicillin, puromycin, tetracycline, kanamycin, bleomycin, streptomycin,
5 hygromycin, neomycin, ZEOCIN™, and the like.

Another element useful in an expression vector is an origin of replication. Replication origins are unique DNA segments that contain multiple short repeated sequences that are recognized by multimeric origin-binding proteins and that play a key role in assembling DNA replication enzymes at the origin site. Suitable origins
10 of replication for use in expression vectors employed herein include *E. coli* oriC, colE1 plasmid origin, and the like.

A further useful element in an expression vector is a multiple cloning site or polylinker. Synthetic DNA encoding a series of restriction endonuclease recognition sites is inserted into a vector, for example, downstream of the promoter
15 element. These sites are engineered for convenient cloning of DNA into the vector at a specific position.

Elements such as the foregoing can be combined to produce expression vectors suitable for use in the methods of the invention. Those of skill in the art will be able to select and combine the elements suitable for use in their particular
20 system in view of the teachings of the present specification.

Provided for is the stable introduction of a large DNA molecule into the cell of an avian. In one particularly useful embodiment, the large DNA molecule is a chromosome. The chromosomes to be introduced into cells of an avian may be referred to herein as “artificial chromosomes”; however, the term “artificial
25 chromosome” is not a limiting term and any useful large DNA molecule or chromosome may be employed in the present invention.

The present invention provides modified chromosomes, which are either isolated chromosomes or artificial chromosomes, which function as useful vectors to shuttle transgenes or gene clusters into the genome. By delivering the modified
30 or artificial chromosome to an isolated recipient cell, the target cell, and progeny thereof, become trisomic or transchromosomic. Typically, an additional or trisomic chromosome will not affect the subsequent development of the recipient

cell and/or an embryo, nor interfere with the reproductive capacity of an adult developed from such cells or embryos. The chromosome also should be stable within chicken cells. An effective method is also required to isolate a population of chromosomes for delivery into chicken embryos or early cells.

5 Chickens that are trisomic for microchromosome 16 have been described (Miller et al, Proc. Natl. Acad. Sci. U.S.A. 93: 3958-3962 (1996); Muscarella et al, J. Cell Biol. 101: 1749-1756 (1985). In these cases, triploidy and trisomy occurred naturally, and illustrate that an extra copy of one or more of the chicken chromosomes is compatible with normal development and reproductive capacity.

10 The transchromosomic avians resulting from the cellular introduction of an artificial chromosome typically will comprise cells which include the normal complement of chromosomes plus at least one additional chromosome. In one embodiment, about 0.001% to 100% of the cells of the avian will include an additional chromosome. In another embodiment, about 0.1% to 100% of the cells
15 of the avian will include an additional chromosome. In another embodiment, about 5% to 100% of the cells of the avian will include an additional chromosome. In another embodiment, about 10% to 100% of the cells of the avian will include an additional chromosome. In another embodiment, about 50% to 100% of the cells of the avian will include an additional chromosome. In one particularly useful
20 embodiment, the additional chromosome is transmitted through the germ-line of the transchromosomic avian and many, for example, most (i.e., more than 50%) of the cells of the offspring avians will include the additional chromosome. The invention contemplates the introduction and propagation of any useful number of chromosomes into the cell(s) of a transgenic avian or isolated avian cells. For
25 example, the invention contemplates one artificial chromosome or two artificial chromosomes or three artificial chromosomes stably incorporated into the genome of the cell(s) of a transchromosomal avian or isolated avian cells.

Any or all tissues of the transchromosomic avian can include the artificial chromosome. In one useful embodiment, one or more cells of the oviduct of the
30 avians include the additional chromosome. For example, tubular gland cells of the oviduct may include the additional chromosome.

A number of artificial chromosomes are useful in the methods of the

invention, including, for instance, a human chromosome modified to work as an artificial chromosome in a heterologous species as described, for example, for mice (Tomizuka et al, Proc. Natl. Acad. Sci. U.S.A. 97: 722-727 (2000); for cattle (Kuroiwa et al, Nat. Biotechnol. 20: 889-894 (2002); a mammalian artificial chromosome used in mice (Co et al, Chromosome Res. 8: 183-191 (2000)).

Examples of large nucleic acid molecules include, but are not limited to, natural chromosomes and fragments thereof, for example, chromosomes (e.g., mammalian chromosomes) and fragments thereof which retain a centromere, artificial chromosome expression systems (satellite DNA-based artificial chromosomes (SATACs); see U.S. Pat. Nos. 6,025,155, issued February 15, 2000 and 6,077,697 issued June 20, 2000, the disclosures of which are incorporated herein in their entirety by reference), mammalian artificial chromosomes (MACs) (e.g., HACs), plant artificial chromosomes, insect artificial chromosomes, avian artificial chromosomes and minichromosomes (see, e.g., U.S. Pat. Nos. 5,712,134 issued January 27, 1998; 5,891,691, issued April 6, 1999; 5,288,625, issued February 22, 1994; 6,743,967 issued June 1, 2004; and U.S. Patent Application Nos. 10/235,119, published June 19, 2003, the disclosure of each of these six patents and the patent application are incorporated herein in their entirety by reference). Also contemplated for use herein are YACs, BACs, bacteriophage-derived artificial chromosomes (BBPACs), cosmid or P1 derived artificial chromosomes (PACs).

As used herein, a large nucleic acid molecule such as artificial chromosomes can stably replicate and segregate alongside endogenous chromosomes in a cell. It has the capacity to act as a gene delivery vehicle by accommodating and expressing foreign genes contained therein. A mammalian artificial chromosome (MAC) refers to chromosomes that have an active mammalian centromere(s). Plant artificial chromosomes, insect artificial chromosomes and avian artificial chromosomes refer to chromosomes that include plant, insect and avian centromeres, respectively. A human artificial chromosome (HAC,) refers to chromosomes that include human centromeres. For exemplary artificial chromosomes, see, e.g., U.S. Pat. Nos. 6,025,155, issued February 15, 2000; 6,077,697, issued June 20, 2000; 5,288,625, issued February 22, 1994;

5,712,134, issued January 27, 1998; 5,695,967, issued December 9, 1997; 5,869,294, issued February 9, 1999; 5,891,691, issued April 6, 1999 and 5,721,118, issued February 24, 1998 and published International PCT application Nos., WO 97/40183, published October 30, 1997 and WO 98/08964, published March 5, 1998, the disclosure of each of these eight patents and two PCT applications are incorporated in their entirety herein by reference.

The large nucleic acid molecules (e.g., chromosomes) can include a single copy of a desired nucleic acid fragment encoding a particular nucleotide sequence, such as a gene of interest, or can carry multiple copies thereof or multiple genes, different heterologous nucleotide sequences or expression cassettes or may encode one or more heterologous transcripts each encoding more than one useful protein product (for example, the transcript(s) may comprise an IRES). Any useful IRES may be employed in the invention. See, for example, US Patent No. 4,937,190, issued January 26, 1990; Nature (1988) 334:320-325; J Virol (1988) 62:3068-3072; Cell (1992) 68:119-131; J Virol (1990) 64:4625-4631; and J Virol (1992) 66:1476-1483, the disclosures of which are incorporated in their entirety herein by reference, which disclose useful IRESs. For example, the nucleic acid molecules can carry 40 or even more copies of genes of interest. The large nucleic acid molecules can be associated with proteins, for example, chromosomal proteins, that typically function to regulate gene expression and/or participate in determining overall structure (e.g., nucleosomes).

Certain useful artificial chromosomes, such as satellite DNA-based artificial chromosomes, can include substantially all neutral non-coding sequences (heterochromatin) except for foreign heterologous, typically gene-encoding, nucleic acid (see U.S. Pat. Nos. 6,025,155, issued February 15, 2000 and 6,077,697, issued June 20, 2000 and International PCT application No. WO 97/40183, published October 30, 1997 and Lindenbaum et al Nucleic Acids Res (2004) vol 32 no. 21 e172, the disclosures of these two patents, the PCT application and the publication are incorporated in their entirety herein by reference). Foreign genes (i.e., nucleotide sequences of interest) contained in these artificial chromosomes can include, but are not limited to, nucleic acid that encodes therapeutically effective substances (e.g., therapeutic proteins such as those disclosed elsewhere herein and

traceable marker proteins (reporter genes), such as fluorescent proteins, such as green, blue or red fluorescent proteins (GFP, BFP and RFP, respectively), other reporter genes, such as beta-galactosidase and proteins that confer drug resistance, such as a gene encoding hygromycin-resistance.

5 In one useful embodiment, the artificial chromosomes employed herein do not interfere with the host cells' processes and can be easily purified by useful purification methods such as large-scale by high-speed flow cytometry (see, for example, de Jong, G, et al. Cytometry 35: 129-33, 1999). Such artificial chromosomes are useful for the production of transchromosomic chickens
10 produced by introduction of the chromosomes into certain cells, for example, the germline cells, of an avian. In one particularly useful embodiment of the present invention, the transchromosomic chickens are produced by microinjection of the chromosomes, for example, cytoplasmic injection of the chromosomes into avian embryos, for example, early stage embryos such as a Stage I embryos, see, for
15 example, US Patent Application No. 10/679,034, filed October 2, 2003, the disclosure of which is incorporated in its entirety herein by reference.

 In one embodiment, heterologous nucleic acid is introduced into an artificial chromosome. Any useful method to introduce the nucleic acid into the chromosome may be employed in the invention. Thereafter, the artificial
20 chromosomes are isolated in a mixture substantially free of other chromosomes or cellular material. For example, artificial chromosomes may be isolated by flow cytometry (e.g., dual laser high-speed flow cytometer as described previously (de Jong, G, et al. Cytometry 35: 129-33, 1999). See, for example, US Patent Application Publication No. 20030113917, published June 19, 2003, the disclosure
25 of which is incorporated in its entirety herein by reference.

 In accordance with the present invention, any useful number of artificial chromosomes may be introduced into an avian cell (e.g., injected), for example, an avian germinal cell such as a cell of an ova, an embryo or a germinal disc of an avian egg. Any useful method of introducing the chromosomes into the avian cell
30 is contemplated for use in the present invention. In addition, the invention contemplates the introduction of any useful number of chromosomes into an avian cell. For example, and without limitation, the invention contemplates the

introduction of 1 to about 1,000,000 chromosomes injected per egg. In one embodiment, 1 to about 100,000 chromosomes are injected per egg. In another embodiment about 5 to about 100,000 artificial chromosomes are injected per egg. For example, about 10 to about 50,000 chromosomes may be injected per egg.

5 In one embodiment, there is a lower hatch rate for eggs injected with more than a certain number of chromosomes. In one embodiment, an injection of over 100,000 chromosomes reduces or brings the hatch rate to zero. In another embodiment, an injection of over 20,000 chromosomes reduces or brings the hatch rate to zero. In another embodiment, an injection of over 5,000 chromosomes
10 reduces or brings the hatch rate to zero. In another embodiment, an injection of over 2,000 chromosomes reduces or brings the hatch rate to zero. For example, an injection of over 1,000 (e.g., 550) chromosomes reduces or brings the hatch rate to zero.

For injection, any useful volume of injection buffer may be used for each
15 injection. For example, about 1 nl to about 1 μ l may be injected. In addition, any useful concentration of chromosomes may be employed in the injection buffer. For example, and without limitation, 1 to about 100,000 chromosomes per microliter may be used. In addition, any useful number of injections may be performed on each egg.

20 In one embodiment, a concentration of 7000-11,500 chromosomes is used per μ l of injection buffer (Monteith, D, et al. Methods Mol Biol 240: 227-242, 2004). In one embodiment, 25-100 nanoliters (nl) of injection buffer is used per injection.

Any useful avian embryos may be employed in the present invention. For
25 example, the embryos may be collected from 24-36 week-old hens (e.g., commercial White Leghorn variety of *G. gallus*). In one embodiment, a germinal disc is injected with the chromosomes. In one embodiment, the embryo donor hens are inseminated weekly using pooled semen from roosters to produce eggs for injection. Any useful method, such as methods known to those skilled in the art,
30 may be employed to collect fertilized eggs.

Cytoplasmic injection of artificial chromosomes can be achieved by employing certain microinjection systems or assemblies. In one particularly useful

embodiment, the microinjection assembly or microinjection system disclosed in US Patent Application No. 09/919,143, filed July 31, 2001 (the '143 application), the disclosure of which is incorporated herein in its entirety, is employed. Use of such a cytoplasmic injection device allows for the precise delivery of chromosomes
5 into the cytoplasm of avian embryos, for example, early stage avian embryos, e.g., Stage I embryos.

Typically, following microinjection, the embryos are transferred to the oviduct of recipient hens utilizing any useful technique, such as that disclosed in Olsen, M and Neher, B. J Exp Zool 109: 355-66, 1948, followed by incubation and
10 hatching of the birds.

Any useful method, such as PCR, may be used to test for the production of transchromosomic avians. Typically, the identification of a transchromosomic offspring is confirmed by fluorescence in-situ hybridization (FISH) and/or DNA analysis such as Southern blot or the like. In one useful embodiment, artificial
15 chromosomes can be used as vectors to introduce large DNA payloads, such as nucleotide sequences to be expressed heterologously in the avian to yield a desired biomolecule, of stably maintained genetic information into transgenic chickens. Production of germ-line transchromosomic avians is confirmed by the production of transchromosomic offspring from the G0 birds.

20 The present invention provides for the introduction of desired nucleotide sequences into a chromosome, the chromosome of which can subsequently be isolated/purified and thereafter introduced into an avian as disclosed herein.

A useful chromosome isolation protocol can comprise the steps of inserting a lac-operator sequence (Robinett et al J. Cell Biol. 135: 1685-1700 (1996) into an
25 isolated chromosome and, optionally, inserting a desired transgene sequence within the same chromosome. In one embodiment, the lac operator region is a concatamer of a plurality of lac operators for the binding of multiple lac repressor molecules. Insertion can be accomplished, for instance, by identifying a region of known nucleotide sequence associated with a particular avian chromosome. A
30 recombinant DNA molecule may be constructed that comprises the identified region, a recombination site such as attB or attP and a lac-operator concatamer. The recombinant molecule is delivered to an isolated avian cell, for example, but

not limited to, chicken DT40 cells that have elevated homologous recombination activity compared to other avian cell lines, whereupon homologous recombination will integrate the heterologous recombination site and the lac-operator concatamer into the targeted chromosome as shown in the schema illustrated in Fig. 20. A tag-
5 polypeptide comprising a label domain and a lac repressor domain is also delivered to the cell, for example, by expression from a suitable expression vector. The nucleotide sequence coding for a GFP-lac-repressor fusion protein (Robinett et al, J. Cell Biol. 135: 1685-1700 (1996)) may be inserted into the same chromosome as the lac-operator insert. The lac repressor sequence, however, can also be within a
10 different chromosome. An inducible promoter may also be used to allow the expression of the GFP-lac-repressor only after chromosome is to be isolated.

Induced expression of the GFP-lac-repressor fusion protein will result in specific binding of the tag fusion polypeptide to the lac-operator sequence for identification and isolation of the genetically modified chromosome. The tagged
15 mitotic chromosome can be isolated using, for instance, flow cytometry as described in de Jong et al Cytometry 35: 129-133 (1999) and Griffin et al Cytogenet. Cell Genet. 87: 278-281 (1999).

A tagged chromosome can also be isolated using microcell technology requiring treatment of cells with the mitotic inhibitor colcemid to induce the
20 formation of micronuclei containing intact isolated chromosomes within the cell. Final separation of the micronuclei is then accomplished by centrifugation in cytochalasin as described by Killary & Fournier in Methods Enzymol. 254: 133-152 (1995). Further purification of microcells containing only the desired tagged chromosome could be done by flow cytometry. It is contemplated, however, that
25 alternative methods to isolate the mitotic chromosomes or microcells, including mechanical isolation or the use of laser scissors and tweezers, and the like.

The present invention envisions the employment of any useful protein-DNA binding or interaction to assist in isolating/purifying chromosomes of the invention. Such other methods in which a desired chromosome can be labeled for
30 purposes of isolation/purification, are well known in the art including but not limited to, steroid receptor (such as the glucocorticoid receptor):site specific response element systems, see, for example, McNally et al, Science 287:1262-

1265; the bacteriophage lambda repressor system; and human homeobox genes. In addition, certain mutant forms of proteins which are employed in these systems (e.g., mutant proteins which bind there substrate with greater affinity than the non-mutant form of the protein) can be particularly useful for chromosome tagging and
5 subsequent isolation/purification of the chromosomes. Furthermore the invention contemplates the use of a selectable marker to identify cells which contain chromosomes comprising an introduced sequence of interest.

For example, as seen in FIG. 25, an artificial chromosome may include a promoter (e.g., SV40) that will express a marker, such as an antibiotic resistant
10 marker (e.g., hygromycin), when a vector (e.g., plasmid) which includes the gene of interest and the marker coding sequence integrates into the chromosome. For example, a useful cell line such as LMTK- containing the chromosome (A) in FIG. 25 is transfected with the vector B by standard methodologies such as lipofection. After introduction of the vector (B) into the artificial chromosome containing cell
15 line, integration occurs, for example, between integration sites such as lambda attB and attP sites, wherein the hygromycin marker is expressed in the cells which contain the recombined artificial chromosome allowing for selection of the cells. For the employment of such integration sites, integrase or an integrase encoding gene is typically also introduced into the cell. In one useful embodiment, a lambda
20 integrase gene is used which produces an integrase protein with a substitution mutation at the glutamine residue at position 174 to a lysine. This mutation removes the requirement for host factors allowing the integrase to function in cell lines.

This is merely an example of a marker system that can be used to select for
25 chromosomes comprising the nucleotide sequence of interest and other similar systems can be readily envisioned by a practitioner of skill in the art. For example, the method of Gygi et al (2002) Nucleic Acids Res. 30: 2790-2799, the disclosure of which is incorporated by reference herein in its entirety, is contemplated for use in the present invention. Briefly, the protocol provides for the use of synthetic
30 polyamide probes to fluorescently label heterochromatic regions on the chromosomes which are then isolated by flow cytometry. The polyamides bind to the minor groove of DNA of the chromosomes in a sequence specific manner

without the need to disrupt the chromosome (e.g., denature the DNA).

Typically, the artificial chromosomes introduced into avians are stably maintained in the avians and are passed to offspring through the germline. In addition, artificial chromosomes can be stably maintained in avian cell lines such as chicken cell line (DT-40).

The invention is also useful for visualizing gene activity in avian cells as is understood by a practitioner of ordinary skill in the art (See, for example, Tsukamoto, et al (2000) *Nature Cell Biology*, 2:871-878).

Most non-viral methods of gene transfer rely on normal mechanisms used by eukaryotic cells for the uptake and intracellular transport of macromolecules. In certain useful embodiments, non-viral gene delivery systems of the present invention rely on endocytic pathways for the uptake of the subject transcriptional regulatory region and operably linked polypeptide-encoding nucleic acid by the targeted cell. Exemplary gene delivery systems of this type include liposomal derived systems, poly-lysine conjugates, and artificial viral envelopes. Modified chromosomes as described above may be delivered to isolated avian embryonic cells for subsequent introduction to an embryo.

In a representative embodiment, a nucleic acid molecule can be entrapped in liposomes bearing positive charges on their surface (e.g., lipofectins) and (optionally) which are tagged with antibodies against cell surface antigens of the target tissue (Mizuno et al, 1992, *NO Shinkei Geka* 20: 547-551; PCT publication WO91/06309, published May 16, 1991; Japanese patent application 1047381, published February 21, 1989; and European patent publication EP-A-43075, published January 6, 1982, all of which are incorporated herein by reference in their entireties).

In similar fashion, the gene delivery system can comprise an antibody or cell surface ligand that is cross-linked with a gene binding agent such as polylysine (see, for example, PCT publications WO93/04701, published March 18, 1993; WO92/22635, published December 23, 1992; WO92/20316, published November 26, 1992; WO92/19749, published November 12, 1992; and WO92/06180, published April 16, 1992, the disclosures of which are incorporated herein by reference in their entireties). It will also be appreciated that effective delivery of

the subject nucleic acid constructs via receptor-mediated endocytosis can be improved using agents which enhance escape of genes from the endosomal structures. For instance, whole adenovirus or fusogenic peptides of the influenza HA gene product can be used as part of the delivery system to induce efficient
5 disruption of DNA-containing endosomes (Mulligan et al, 1993, Science 260:926-932; Wagner et al, 1992, Proc. Natl. Acad. Sci. 89:7934-7938; and Christiano et al, 1993, Proc. Natl. Acad. Sci. 90:2122-2126, all of which are incorporated herein by reference in their entireties). It is further contemplated that a recombinant nucleic acid molecule of the present invention may be delivered to a target host cell by
10 other non-viral methods including by gene gun, microinjection, sperm-mediated transfer, or the like.

In one embodiment of the invention, an expression vector that comprises a recombination site, such as an attB site, and a region encoding a polypeptide deposited into an egg white are delivered to oviduct cells by in vivo
15 electroporation. In this method, the luminal surface of an avian oviduct is surgically exposed. A buffered solution of the expression vector and a source of integrase activity such as a second expression vector expressing integrase (for example, pCMV-int) is deposited on the luminal surface. Electroporation electrodes are then positioned on either side of the oviduct wall, the luminal
20 electrode contacting the expression vector solution. After electroporation, the surgical incisions are closed. The electroporation will deliver the expression vectors to some, if not all, treated recipient oviduct cells to create a tissue-specific chimeric animal. Expression of the integrase allows for the integration of the heterologous polynucleotide into the genome of recipient oviduct cells. While this
25 method may be used with any bird, a useful recipient is a chicken due to the size of the oviduct. Also useful is a transgenic bird that has a transgenic attP recombinant site in the nuclear genomes of recipient oviduct cells, thus increasing the efficiency of integration of the expression vector.

The attB/P integrase system is useful in the in vivo electroporation method
30 to allow the formation of stable genetically transformed oviduct cells that otherwise progressively lose the heterologous expression vector.

The stably modified oviduct cells will express the heterologous

polynucleotide and deposit the resulting polypeptide into the egg white of a laid egg. For this purpose, the expression vector will further comprise an oviduct-specific promoter such as ovalbumin or ovomucoid operably linked to the desired heterologous polynucleotide.

5 Another aspect of the invention is the generation of a trisomic or transchromosomic avian cell comprising a genetically modified extra chromosome. The extra chromosome may be an artificial chromosome or an isolated avian chromosome that has been genetically modified. Introduction of the extra chromosome to an avian cell will generate a trisomic or transchromosomic cell
10 with $2n+1$ chromosomes, where n is the haploid number of chromosomes of a normal avian cell.

 Delivery of an isolated chromosome into an isolated avian cell or embryo can be accomplished in several ways. Isolated mitotic chromosomes or a micronucleus containing an interphase chromosome can be injected into early stage
15 I embryos by cytoplasmic injection. The injected zygote would then be surgically transferred to a recipient hen for the production and laying of a hard shell egg. This hard shell egg would then be incubated until hatching of a chick.

 In one embodiment, isolated microcells which contain the artificial chromosome can be fused to primordial germ cells (PGCs) isolated from the blood
20 stream of late stage 15 embryos as described by Killary & Fournier in *Methods Enzymol.* 254: 133-152 (1995). The PGC/microcell hybrids can then be transplanted into the blood stream of a recipient embryo to produce germline chimeric chickens. (See Naito et al, *Mol. Reprod. Dev.* 39: 153-161 (1994)). The manipulated eggs would then incubated until hatching of the bird.

25 Blastodermal cells isolated from stage X embryos can be transfected with isolated mitotic chromosomes. Following in vitro transfection, the cells are transplanted back into stage X embryos as described, for example, in Etches et al, *Poult. Sci.*, 72: 882-889 (1993), and the manipulated eggs are incubated to hatching.

30 Stage X blastodermal cells can also be fused with isolated microcells and then transplanted back into to stage X embryos or fused to somatic cells to be used as nuclear donors for nuclear transfer as described by Kuroiwa et al, *Nat.*

Biotechnol. 20: 889-894 (2002).

Chromosomal vectors, as described above, may be delivered to a recipient avian cell by, for example, microinjection, liposomal delivery or microcell fusion.

5 In the methods of the invention, a site-specific integrase is introduced into an avian cell whose genome is to be modified. Methods of introducing functional proteins into cells are well known in the art. Introduction of purified integrase protein can ensure a transient presence of the protein and its activity. Thus, the lack of permanence associated with most expression vectors is not expected to be detrimental.

10 The integrase used in the practice of the present invention can be introduced into a target cell before, concurrently with, or after the introduction of a site-specific vector. The integrase can be directly introduced into a cell as a protein, for example, by using liposomes, coated particles, or microinjection, or into the blastodermal layer of an early stage avian embryo by microinjection. A source of
15 the integrase can also be delivered to an avian cell by introducing to the cell an mRNA encoding the integrase and which can be expressed in the recipient cell as an integrase polypeptide. Alternately, a DNA molecule encoding the integrase can be introduced into the cell using a suitable expression vector.

The present invention provides novel nucleic acid vectors and methods of
20 use that allow integrases, such as phiC31 integrase, to efficiently integrate a heterologous nucleic acid into a vertebrate animal genome, for example, an avian genome. A novel finding is that the phiC31 integrase is remarkably efficient in avian cells and increases the rate of integration of heterologous nucleic acid at least 30-fold over that of random integration. Furthermore, the phiC31 integrase works
25 equally well at 37°C and 41°C, indicating that it will function in the environment of the developing avian embryo, as shown in Example 1.

It is important to note that the present invention is not bound by any mechanism or theory of operation. For example, the mechanism by which
30 integrase, or any other substance described herein, facilitates transgenesis is unimportant. Integrase, for example, may facilitate transgenesis by mediating the integration of DNA into the genome of a recipient cell or integrase may facilitate transgenesis by facilitating the entry of the DNA into the cell or integrase may

facilitate transgenesis by some other mechanism.

The site-specific vector components described above are useful in the construction of expression cassettes containing sequences encoding an integrase. One integrase-expressing vector useful in the methods of the invention is pCMV-
5 C31int (SEQ ID NO: 1 as shown in Fig. 9) where the phiC31 integrase is encoded by a region under the expression control of the strong CMV promoter. Another useful promoter is the RSV promoter as used in SEQ ID NO: 9 shown in Fig. 17. Expression of the integrase is typically desired to be transient. Accordingly, vectors providing transient expression of the integrase are useful. However,
10 expression of the integrase can be regulated in other ways, for example, by placing the expression of the integrase under the control of a regulatable promoter (i.e., a promoter whose expression can be selectively induced or repressed).

Delivery of the nucleic acids introduced into cells, for example, embryonic cells (e.g., avian cells), using methods of the invention may also be enhanced by
15 mixing the nucleic acid to be introduced with a nuclear localization signal (NLS) peptide prior to introduction, for example, microinjection, of the nucleic acid. Nuclear localization signal (NLS) sequences are a class of short amino acid sequences which may be exploited for cellular import of linked cargo into a nucleus. The present invention envisions the use of any useful NLS peptide,
20 including but not limited to, the NLS peptide of SV40 virus T-antigen.

An NLS of the invention is an amino acid sequence which mediates nuclear transport into the nucleus, wherein deletion of the NLS reduces transport into the nucleus. In certain embodiments, an NLS is a cationic peptide, for example, a highly cationic peptide. The present invention includes the use of any NLS
25 sequence, including but not limited to, SV40 virus T-antigen. NLSs known in the art include, but are not limited to those discussed in Cokol et al, 2000, EMBO Reports, 1(5):411-415, Bouliskas, T., 1993, Crit. Rev. Eukaryot. Gene Expr., 3:193-227, Collas, P. et al, 1996, Transgenic Research, 5: 451-458, Collas and Alestrom, 1997, Biochem. Cell Biol. 75: 633-640, Collas and Alestrom, 1998, Transgenic
30 Research, 7: 303-309, Collas and Alestrom, Mol. Reprod. Devel., 1996, 45:431-438. The disclosure of each of these references is incorporated by reference herein in its entirety.

Not to be bound by any mechanism of operation, DNA is protected and hence stabilized by cationic polymers. The stability of DNA molecules in the cytoplasm of cells may be increased by mixing the DNA to be introduced, for example, microinjected with cationic polymers (for example, branched cationic polymers), such as polyethylenimine (PEI), polylysine, DEAE-dextran, starburst dendrimers, starburst polyamidoamine dendrimers, and other materials that package and condense the DNA molecules (Kukowska-Latallo et al, 1996, Proc. Natl. Acad. Sci. USA 93:4897-4902).

Once the DNA molecules are delivered to the cytoplasm of cells, they migrate into the cell's endocytotic vesicles. Furthermore, migration into the cell's endosome is followed by fast inactivation of DNA within the endolysosomal compartment in transfected or injected cells, both in vitro and in vivo (Godbey, W, et al 1999, Proc Natl Acad Sci U S A 96: 5177-5181; and Lechardeur, D, et al 1999, Gene Ther 6: 482-497; and references cited therein). Accordingly, in certain embodiments, DNA uptake is enhanced by the receptor-mediated endocytosis pathway using transferrin-polylysine conjugates or adenoviral-mediated vesicle disruption to effect the release of DNA from endosomes. However, the invention is not limited to this or any other theory or mechanism of operation referred to herein.

Buffering the endosomal pH using endosomal-scaping elements also protects DNA from degradation (Kircheis, R, et al 2001, Adv Drug Deliv Rev 53: 341-358 ; Boussif, O, et al 1995, Proc Natl Acad Sci U S A 92: 7297-7301; and Pollard, H, et al 1998, J Biol Chem 273: 7507-7511; and references cited therein). Thus, in certain embodiments, DNA complexes are delivered with polycations or cationic polymers that possess substantial buffering capacity below physiological pH, such as polyethylenimine, lipopolyamines and polyamidoamine polymers. In certain embodiments, DNA condensing compounds, such as the ones described above, are combined with viruses (Curiel, D, et al Proc Natl Acad Sci U S A 88: 8850-8854, 1991; Wagner, E, et al Proc Natl Acad Sci U S A 89: 6099-6103, 1992 and Cotten, M, et al, 1992, Proc Natl Acad Sci U S A 89: 6094-6098), viral peptides (Wagner, E, et al 1992, Proc Natl Acad Sci U S A 89: 7934-7938; Plank, C, et al 1994, J Biol Chem 269: 12918-12924) and subunits of toxins (Uherek, C,

et al, 1998, J Biol Chem 273: 8835-48). These materials significantly enhance the release of DNA from endosomes. In certain embodiments, viruses, viral peptides, toxins or subunits of toxins may be coupled to DNA/polylysine complexes via biochemical means or specifically by a streptavidin-biotin bridge (Wagner et al, 5 1992, Proc. Natl. Acad. Sci. USA 89:6099-6103; Plank et al, 1994, J. Biol Chem. 269(17):12918-12924). In other certain embodiments, the virus that is complexed with the DNA may be adenovirus, retrovirus, vaccinia virus, or parvovirus. The viruses may be linked to PEI or another cationic polymer associated with the nucleic acid. In certain embodiments, the virus may be alphavirus, 10 orthomyxovirus, or picornavirus. In certain embodiments, the virus is defective or chemically inactivated. The virus may be inactivated by short-wave UV radiation or the DNA intercalator psoralen plus long-wave UV. The adenovirus may be coupled to polylysine, either enzymatically through the action of transglutaminase or biochemically by biotinylating adenovirus and streptavidinylating the polylysine moiety. Transferrin may also be useful in combination with cationic polymers, 15 adenoviruses and/or other materials disclosed herein to produce transgenic avians. For example, DNA complexes containing PEI, PEI-modified transferrin, and PEI-bound influenza peptides may be used to enhance transgenic avian production.

In other certain embodiments, complexes containing plasmid DNA, 20 transferrin-PEI conjugates, and PEI-conjugated peptides derived from the N-terminal sequence of the influenza virus hemagglutinin subunit HA-2 may be used to produce transgenic chickens. In certain embodiments, the PEI-conjugated peptide may be an amino-terminal amino acid sequence of influenza virus hemagglutinin which may be elongated by an amphipathic helix or by carboxyl- 25 terminal dimerization.

The present invention provides for methods of dispersing or distributing nucleic acid in a cell, for example, in an avian cell. The avian cell may be, for example, and without limitation, a cell of a stage I avian embryo, a cell of a stage II avian embryo, a cell of a stage III avian embryo, a cell of a stage IV avian embryo, 30 a cell of a stage V avian embryo, a cell of a stage VI avian embryo, a cell of a stage VII avian embryo, a cell of a stage VIII avian embryo, a cell of a stage IX avian embryo, a cell of a stage X avian embryo, a cell of a stage XI avian embryo or a

cell of a stage XII avian embryo. In one particularly useful embodiment, the avian cell is a cell of a stage X avian embryo.

In one aspect of the present invention, cationic polymers are useful to distribute, for example, homogeneously distribute, nucleic acid introduced into a cell, for example, an embryonic avian cell. The present invention contemplates the use of cationic polymers including, but not limited to, those disclosed herein.

However, substances other than cationic polymers also capable of distributing or dispersing nucleic acids in a cell are included within the scope of the present invention.

The concentration of cationic polymer used is not critical though, in one useful embodiment, enough cationic polymer is present to coat the nucleic acid to be introduced into the avian cell. The cationic polymer may be present in an aqueous mixture with the nucleic acid to be introduced into the cell at a concentration in a range of an amount equal to about the weight of the nucleic acid to a concentration wherein the solution is saturated with cationic polymer. In one useful embodiment, the cationic polymer is present in an amount in a range of about 0.01% to about 50 %, for example, about 0.1% to about 20% (e.g., about 5%). The molecular weights of the cationic polymers can range from a molecular weight of about 1,000 to a molecular weight of about 1,000,000. In one embodiment, the molecular weight of the cationic polymers range from about 5,000 to about 100,000 for example, about 20,000 to about 30,000.

In one particularly useful aspect of the invention, procedures that are effective to facilitate the production of a transgenic avian may be combined to provide for an enhanced production of a transgenic avian wherein the enhanced production is an improved production of a transgenic avian relative to the production of a transgenic avian by only one of the procedures employed in the combination. For example, one or more of integrase activity, NLS, cationic polymer or other technique useful to enhance transgenic avian production disclosed herein can be used in the same procedure to provide for an enhanced production of transgenic avians relative to an identical procedure which does not employ all of the same techniques useful to enhance transgenic avian production.

Another aspect of the present invention is a vertebrate animal cell which has been genetically modified with a transgene vector according to the present invention and as described herein. For example, in one embodiment, the transformed cell can be a chicken early stage blastodermal cell or a genetically
5 transformed cell line, including a sustainable cell line. The transfected cell according to the present invention may comprise a transgene stably integrated into the nuclear genome of the recipient cell, thereby replicating with the cell so that each progeny cell receives a copy of the transfected nucleic acid. A particularly useful cell line for the delivery and integration of a transgene comprises a
10 heterologous attP site that can increase the efficiency of integration of a polynucleotide by phiC31 integrase and, optionally, a region for expressing the integrase.

A retroviral vector can be used to deliver a recombination site such as an att site into the cellular genomes, such as avian genomes, since an attP or attB site is
15 less than 300 bp. For example, the attP site can be inserted into the NLB retroviral vector, which is based on the avian leukosis virus genome. A lentiviral vector is a particularly suitable vector because lentiviral vectors can transduce non-dividing cells, so that a higher percentage of cells will have an integrated attP site.

The lacZ region of NLB is replaced by the attP sequence. A producer cell
20 line would be created by transformation of, for example, the Isolde cell line capable of producing a packaged recombinant NLB-attP virus pseudo-typed with the envA envelope protein. Supernatant from the Isolde NLB-attP line is concentrated by centrifugation to produce high titer preparations of the retroviral vector that can then be used to deliver the attP site to the genome of a cell, for example, as
25 described in Example 9 below.

In one embodiment, an attP-containing line of transgenic birds are a source of attP transgenic embryos and embryonic cells. Fertile zygotes and oocytes bearing a heterologous attP site in either the maternal, paternal, or both, genomes can be used for transgenic insertion of a desired heterologous polynucleotide. A
30 transgene vector bearing an attB site, for example, would be injected into the cytoplasm along with either an integrase expression plasmid, mRNA encoding the integrase or the purified integrase protein. The oocyte or zygote is then cultured to

hatch by ex ovo methods or reintroduced into a recipient hen such that the hen lays a hard shell egg the next day containing the injected egg.

In another example, fertile stage I to XII embryos, for example, stage VII to XII embryos, hemizygous or homozygous for the heterologous integration site, for example, the attP sequence, may be used as a source of blastodermal cells. The cells are harvested and then transfected with a transgene vector bearing a second recombination site, such as an attB site, plus a nucleotide sequence of interest along with a source of integrase. The transfected cells are then injected into the subgerminal cavity of windowed fertile eggs. The chicks that hatch will bear the nucleotide sequence of interest and the second integration site integrated into the attP site in a percentage of their somatic and germ cells. To obtain fully transgenic birds, chicks are raised to sexual maturity and those that are positive for the transgene in their semen are bred to non-transgenic mates. As disclosed herein, in certain embodiments, the cells of the invention, e.g., embryos, may include an integrase which specifically recognizes recombination sites and which is introduced into cells containing a nucleic acid construct of the invention under conditions such that the nucleic acid sequence(s) of interest will be inserted into the nuclear genome. Methods for introducing such an integrase into a cell are described herein. In some embodiments, the site-specific integrase is introduced into the cell as a polypeptide. In alternative embodiments, the site-specific integrase is introduced into the transgenic cell as a polynucleotide encoding the integrase, such as an expression cassette optionally carried on a transient expression vector, and comprising a polynucleotide encoding the recombinase.

In one embodiment, the invention is directed to methods of using a vector for site-specific integration of a heterologous nucleotide sequence into the genome of a cell, the vector comprising a circular backbone vector, a polynucleotide of interest operably linked to a promoter, and a first recombination site, wherein the genome of the cell comprises a second recombination site and recombination between the first and second recombination sites is facilitated by an integrase. In certain embodiments, the integrase facilitates recombination between a bacterial genomic recombination site (attB) and a phage genomic recombination site (attP).

In another embodiment, the invention is directed to a cell having a transformed genome comprising an integrated heterologous polynucleotide of interest whose integration, mediated by an integrase, was into a recombination site native to the cell genome and the integration created a recombination-product site comprising the polynucleotide sequence. In yet another embodiment, integration of the polynucleotide was into a recombination site not native to the cell genome, but instead into a heterologous recombination site engineered into the cell genome.

In further embodiments, the invention is directed to transgenic vertebrate animals, such as transgenic birds, comprising a modified cell and progeny thereof as described above, as well as methods of producing the same.

For example, cells genetically modified to carry a heterologous attB or attP site by the methods of the present invention can be maintained under conditions that, for example, keep them alive but do not promote growth and/or cause the cells to differentiate or dedifferentiate. Cell culture conditions may be permissive for the action of the integrase in the cells, although regulation of the activity of the integrase may also be modulated by culture conditions (e.g., raising or lowering the temperature at which the cells are cultured).

One aspect of the invention are methods for generating a genetically modified cell for example, an avian cell, and progeny thereof, using a tagged chromosome. The methods may include providing an isolated modified chromosome comprising a lac operator region and a first recombination site, delivering the modified chromosome to a avian cell, thereby generating a trisomic or transchromosomic avian cell, delivering to the avian cell a source of a tagged polypeptide comprising a fluorescent domain and a lac repressor domain, delivering a source of integrase activity to the avian cell, delivering a polynucleotide comprising a second recombination site and a region encoding a polypeptide to the avian cell, maintaining the avian cell under conditions suitable for the integrase to mediate recombination between the first and second recombination sites, thereby integrating the polynucleotide into the modified chromosome and generating a genetically modified avian cell, expressing the tag polypeptide by the avian cell, allowing the tag polypeptide to bind to the modified chromosome so as to label the modified chromosome, and isolating the modified

chromosome by selecting modified chromosomes having a tag polypeptide bound thereto.

In one embodiment of the invention, the second avian cell is selected from the group consisting of a stage VII-XII blastodermal cell, a stage I embryo, a stage
5 X embryo; an isolated primordial germ cell, an isolated non-embryonic cell, and an oviduct cell.

In various embodiments, the isolated modified chromosome is an avian chromosome or an artificial chromosome.

In other embodiments of the invention, the step of providing an isolated
10 modified chromosome comprising a lac operator region and a first recombination site comprises the steps of generating a trisomic or transchromosomic avian cell by delivering to an isolated avian cell an isolated chromosome and a polynucleotide comprising a lac operator and a second recombination site, maintaining the trisomic or transchromosomic cell under conditions whereby the heterologous
15 polynucleotide is integrated into the chromosome by homologous recombination, delivering to the avian cell a source of a tag polypeptide to label the chromosome, and isolating the labeled chromosome.

In one embodiment of the invention, the lac operator region is a concatamer of lac operators. In other embodiments of the invention, the tag polypeptide is
20 expressed from an expression vector.

In one embodiment of the invention, the tag polypeptide is microinjected into the cell. In various embodiments of the invention, the method of delivery of a chromosome to an avian cell is selected from the group consisting of liposome delivery, microinjection, microcell, electroporation and gene gun delivery, or a
25 combination thereof.

In embodiments of the invention, the fluorescent domain of the tag polypeptide is GFP.

In one embodiment of the invention, the method further comprises the step of delivering the second avian cell to an avian embryo. The embryo may be
30 maintained under conditions suitable for hatching as a chick.

In one embodiment of the invention, the second avian cell is maintained under conditions suitable for the proliferation of the cell, and progeny thereof.

In various embodiments of the invention, the source of integrase activity is delivered to a first avian cell as a polypeptide or expressed from a polynucleotide, said polynucleotide being selected from an mRNA and an expression vector.

5 In one embodiment of the invention, the tag polypeptide activity is delivered to the avian cell as a polypeptide or expressed from a polynucleotide operably linked to a promoter. In another embodiment of the invention, the promoter is an inducible promoter. In yet another embodiment of the invention, the integrase is phiC31 integrase and in various embodiments of the invention, the first and second recombination sites are selected from an attB and an attP site, but
10 wherein the first and second sites are not identical.

Other aspects of the present invention include methods of expressing a heterologous polypeptide in vertebrate cells by stably transfecting cells using site-specific integrase-mediation and a recombinant nucleic acid molecule, as described herein, and culturing the transfected cells under conditions suitable for expression
15 of the heterologous polypeptide. In addition, the present invention includes methods of expressing a heterologous polypeptide in a transgenic vertebrate animal by producing a transgenic vertebrate animal using methods known in the field or described herein in combination with using site-specific integration of nucleic acid molecules as described herein, and exposing the animal to conditions suitable for
20 expression of the heterologous polypeptide.

The protein of the present invention may be produced in purified form by any known conventional techniques. For example, in the case of heterologous protein production in eggs, the egg white may be homogenized and centrifuged. The supernatant may then be subjected to sequential ammonium sulfate
25 precipitation and heat treatment. The fraction containing the protein of the present invention is subjected to gel filtration in an appropriately sized dextran or polyacrylamide column to separate the proteins. If necessary, the protein fraction may be further purified by HPLC or other methods well known in the art of protein purification.

30 The methods of the invention are useful for expressing nucleic acid sequences that are optimized for expression in the host cells and which encode desired polypeptides or derivatives and fragments thereof. Derivatives include, for

instance, polypeptides with conservative amino acid replacements, that is, those within a family of amino acids that are related in their side chains (commonly known as acidic, basic, nonpolar, and uncharged polar amino acids). Phenylalanine, tryptophan, and tyrosine are sometimes classified jointly as aromatic amino acids and other groupings are known in the art (see, for example, 5 "Biochemistry", 2nd ed, L. Stryer, ed., W.H. Freeman & Co., 1981). Peptides in which more than one replacement has taken place can readily be tested for activity in the same manner as derivatives with a single replacement, using conventional polypeptide activity assays (e.g. for enzymatic or ligand binding activities).

10 Regarding codon optimization, if the recombinant nucleic acid molecules are transfected into a recipient chicken cell, the sequence of the nucleic acid insert to be expressed can be optimized for chicken codon usage. This may be determined from the codon usage of at least one, or more than one, protein expressed in a chicken cell according to well known principles. For example, in 15 the chicken the codon usage could be determined from the nucleic acid sequences encoding the proteins such as lysozyme, ovalbumin, ovomucin and ovotransferrin of chicken. Optimization of the sequence for codon usage can elevate the level of translation in avian eggs.

20 The present invention provides methods for the production of a protein by cells comprising the steps of maintaining a cell, transfecting with a first expression vector and, optionally, a second expression vector, under conditions suitable for proliferation and/or gene expression and such that an integrase will mediate site specific recombination at att sites. The expression vectors may each have a transcription unit comprising a nucleotide sequence encoding a heterologous 25 polypeptide, wherein one polypeptide is an integrase, a transcription promoter, and a transcriptional terminator. The cells may then be maintained under conditions for the expression and production of the desired heterologous polypeptide(s).

30 The present invention further relates to methods for gene expression by cells, such as avian cells, from nucleic acid vectors, and transgenes derived therefrom, that include more than one polypeptide-encoding region wherein, for example, a first polypeptide-encoding region can be operatively linked to an avian promoter and a second polypeptide-encoding region is operatively linked to an

Internal Ribosome Entry Sequence (IRES). It is contemplated that the first polypeptide-encoding region, the IRES and the second polypeptide-encoding region of a recombinant DNA of the present invention may be arranged linearly, with the IRES operably positioned immediately 5' of the second polypeptide-
5 encoding region. This nucleic acid construct can be used for the production of certain proteins in vertebrate animals or in their cells. For example, when inserted into the genome of an avian cell or a bird and expressed therein, will generate individual polypeptides that may be post-translationally modified and combined in the white of a hard shell bird egg. Alternatively, the expressed polypeptides may
10 be isolated from an avian egg and combined in vitro.

The invention, therefore, includes methods for producing multimeric proteins including immunoglobulins, such as antibodies, and antigen binding fragments thereof. Thus, in one embodiment of the present invention, the multimeric protein is an immunoglobulin, wherein the first and second
15 heterologous polypeptides are immunoglobulin heavy and light chains respectively. Illustrative examples of this and other aspects of the present invention for the production of heterologous multimeric polypeptides in avian cells are fully disclosed in U.S. Patent Application No. 09/877,374, filed June 8, 2001, and U.S. Patent Application No. 10/251,364, filed September 18, 2002, both of which are
20 incorporated herein by reference in their entirety.

Accordingly, the invention further provides immunoglobulin and other multimeric proteins that have been produced by transgenic vertebrates including avians of the invention.

In various embodiments, an immunoglobulin polypeptide encoded by the
25 transcriptional unit of at least one expression vector may be an immunoglobulin heavy chain polypeptide comprising a variable region or a variant thereof, and may further comprise a D region, a J region, a C region, or a combination thereof. An immunoglobulin polypeptide encoded by an expression vector may also be an immunoglobulin light chain polypeptide comprising a variable region or a variant
30 thereof, and may further comprise a J region and a C region. The present invention also contemplates multiple immunoglobulin regions that are derived from the same animal species, or a mixture of species including, but not only, human, mouse, rat,

rabbit and chicken. In certain embodiments, the antibodies are human or humanized.

In other embodiments, the immunoglobulin polypeptide encoded by at least one expression vector comprises an immunoglobulin heavy chain variable region,
 5 an immunoglobulin light chain variable region, and a linker peptide thereby forming a single-chain antibody capable of selectively binding an antigen.

Examples of therapeutic antibodies that may be produced in methods of the invention include but are not limited to HERCEPTINTM (Trastuzumab) (Genentech, CA) which is a humanized anti-HER2 monoclonal antibody for the
 10 treatment of patients with metastatic breast cancer; REOPROTM (abciximab) (Centocor) which is an anti-glycoprotein IIb/IIIa receptor on the platelets for the prevention of clot formation; ZENAPAXTM (daclizumab) (Roche Pharmaceuticals, Switzerland) which is an immunosuppressive, humanized anti-CD25 monoclonal antibody for the prevention of acute renal allograft rejection; PANOREXTM which
 15 is a murine anti-I7-IA cell surface antigen IgG2a antibody (Glaxo Wellcome/Centocor); BEC2 which is a murine anti-idiotypic (GD3 epitope) IgG antibody (ImClone System); IMC-C225 which is a chimeric anti-EGFR IgG antibody (ImClone System); VITAXINTM which is a humanized anti- α V β 3 integrin antibody (Applied Molecular Evolution/MedImmune); Campath 1H/LDP-03 which
 20 is a humanized anti CD52 IgG1 antibody (Leukosite); Smart M195 which is a humanized anti-CD33 IgG antibody (Protein Design Lab/Kanebo); RITUXANTM which is a chimeric anti-CD20 IgG1 antibody (IDEC Pharm/Genentech, Roche/Zettyaku); LYMPHOCIDETM which is a humanized anti-CD22 IgG antibody (Immunomedics); ICM3 is a humanized anti-ICAM3 antibody (ICOS Pharm); IDEC-114 is a primate anti-CD80 antibody (IDEC Pharm/Mitsubishi); ZEVALINTM is a radiolabelled murine anti-CD20 antibody (IDEC/Schering AG); IDEC-131 is a humanized anti-CD40L antibody (IDEC/Eisai); IDEC-151 is a primatized anti-CD4 antibody (IDEC); IDEC-152 is a primatized anti-CD23 antibody (IDEC/Seikagaku); SMART anti-CD3 is a humanized anti-CD3 IgG
 25 (Protein Design Lab); 5G1.1 is a humanized anti-complement factor 5 (CS) antibody (Alexion Pharm); D2E7 is a humanized anti-TNF- α antibody (CATIBASF); CDP870 is a humanized anti-TNF- α Fab fragment (Celltech);

IDEC-151 is a primatized anti-CD4 IgG1 antibody (IDEC Pharm/SmithKline Beecham); MDX-CD4 is a human anti-CD4 IgG antibody (Medarex/Eisai/Genmab); CDP571 is a humanized anti-TNF- α IgG4 antibody (Celltech); LDP-02 is a humanized anti- $\alpha 4\beta 7$ antibody (LeukoSite/Genentech);
5 OrthoClone OKT4A is a humanized anti-CD4 IgG antibody (Ortho Biotech); ANTOVA™ is a humanized anti-CD40L IgG antibody (Biogen); ANTEGREN™ is a humanized anti-VLA-4 IgG antibody (Elan); and CAT-152 is a human anti-TGF- β_2 antibody (Cambridge Ab Tech).

The invention can be used to express, in large yields and at low cost, a wide
10 range of desired proteins including those used as human and animal pharmaceuticals, diagnostics, and livestock feed additives. Proteins such as growth hormones, cytokines, structural proteins and enzymes including human growth hormone, interferon, lysozyme, and β -casein are examples of proteins which are desirably expressed in the oviduct and deposited in eggs according to the invention.
15 Other possible proteins to be produced include, but are not limited to, albumin, α -1 antitrypsin, antithrombin III, collagen, factors VIII, IX, X (and the like), fibrinogen, hyaluronic acid, insulin, lactoferrin, protein C, erythropoietin (EPO), granulocyte colony-stimulating factor (G-CSF), granulocyte macrophage colony-stimulating factor (GM-CSF), tissue-type plasminogen activator (tPA), feed additive enzymes,
20 somatotropin, and chymotrypsin. Immunoglobulins (shown, for example in Example 10 below) and genetically engineered antibodies, including immunotoxins which bind to surface antigens on human tumor cells and destroy them, can also be expressed for use as pharmaceuticals or diagnostics.

Other specific examples of therapeutic substances which are contemplated
25 for production as disclosed herein include, with out limitation, Factor VIII (e.g., Recombinate®, Bioclone®, Kogenate®, Helixate® (Centeon), B-domain deleted Factor VIII (e.g., ReFacto®), Factor VIIa (e.g., NovoSeven®), Factor IX (e.g., Benefix®), anticoagulant; recombinant hirudin (e.g., Revasc®, Refludan®) Alteplase, tPA (e.g., Activase®), Reteplase, tPA, tPA – 3 of 5 domains deleted,
30 Ecokinase®, Retavase®, Rapilysin®, insulin (e.g., Humulin®, Novolin®, Insuman®) insulin lispro (e.g., Humalog®), Bio Lysprol, Liprolog®), insulin Aspart, iNovoRapid®, insulin glargine, long-acting insulin analog (e.g., Lantus®),

rhGH (e.g., Protropin ®, Humatrope®, Nutropin®, BioTropin®, Genotropin®, Norditropin®, Saizen®, Serostim®), glucagons (e.g., Glucagen®), TSH (e.g., Thyrogen®, Gonal F®, Puregon®), follitropin-beta FSH (e.g., Follistim®), EPO (e.g., Epogen®, Procrit®, Neorecormon®), GM-CSF (e.g., Leukine®,

5 Neupogen®), PDGH (e.g., Regranex®), IFN alpa2a (e.g., Roferon A®), INF-apha (e.g., Infergen®), IFN alpa2b (e.g., Intron A®, Alfatronol®, Virtron®), ribavirin & INF-alpha 2b (e.g., Robetron®) INF-beta 1b, differs from h protein by C17 to S (e.g., Betaferon®), IFN-beta 1a (e.g., Avonex®, Rebif®), IFN-gamma1b (e.g., Actimmune®), IL-2 (e.g., Proleukin®) rIL-11 (e.g., Neumega®), rHBsAg (e.g.,

10 Recombivax®), Combination vaccine containing HBsAgn as one component (e.g., Comvax®, Tritarix®, Twinrix®, Primavax®, Procomax®), OspA, a lipoprotein found on the surface of B burgoeri (e.g., Lymerix®), murine MAb directed against t-lymphocyte antigen CD3 (e.g., Orthoclone OKT3®), murine MAb directed against TAG-72, tumor-associated glycoprotein (e.g., OncoScint CR/OV®), FAb

15 fragments derived from chimeric MAb, directed against platelet surface receptor GPII(b)/III(a) (e.g., ReoPro®), murine MAb fragment directed against tumor-associated antigen CA125 (e.g., Indimacis®), murine MAb fragment directed against human carcinoembryonic antigen, CEA (e.g., CEA-scan®), murine MAb fragment directed against human cardiac myosin (e.g., MyoScint®), murine MAb

20 fragment directed against tumor surface antigen PSMA (e.g., ProstaScint®), murine MAb fragments (FAb/FAb2 mix) directed against HMW-MAA (e.g., Tacnemab®), murine MAb fragment (FAb) directed against carcinoma-associated antigen (e.g.,Verluma®), MAb fragments (FAb) directed against NCA 90, a surface granulocyte nonspecific cross reacting antigen (e.g., LeukoScan®),

25 chimeric MAb directed against CD20 antigen found on surface of B lymphocytes (e.g., Rituxan®), humanized MAb directed against the alpha chain of the IL2 receptor (e.g., Zenapax®), chimeric MAb directed against the alpha chain of the IL2 receptor (e.g., Simulect®), chimeric MAb directed against TNF-alpha (e.g., Remicade®), humanized MAb directed against an epitope on the surface of

30 respiratory syncytial virus (e.g., Synagis®), humanized MAb directed against HER 2, i.e., human epidermal growth factor receptor 2 (e.g., Herceptin®), human MAb directed against cytokeratin tumor-associated antigen (e.g., Humaspect®), anti-

CTLA4, chimeric MAb directed against CD 20 surface antigen of B lymphocytes (e.g., Mabthera®), dornase-alpha DNase (e.g., Pulmozyme®), beta glucocerebrosidase (e.g., Cerezyme®), TNF-alpha (e.g., Beromun®), IL-2-diphtheria toxin fusion protein that targets cells displaying a surface IL-2 receptor (e.g., Ontak®), TNFR-IgG fragment fusion protein (e.g., Enbrel®), Laronidase, Recombinant DNA enzyme, (e.g., Aldurazyme®), Alefacept, Amevive®, Darbepoetin alfa (Colony stimulating factor) (e.g., Aranesp®), Tositumomab and iodine 1 131 tositumomab, murine MAb, Bexxar®, Alemtuzumab, Campath®, Rasburicase, Elitck®, Agalsidase beta, Fabrazyme®, FluMist®, Teriparatide, Parathyroid hormone derivative (e.g., Forteo®), Enfuvirtide Fuzeon®, Adalimumab (IgG1) (e.g., Humira®), Anakinra, Biological modifier (e.g., Kineret®), nesiritide, Human B-type natriuretic peptide (hBNP) (e.g., Natrecor®), Pegfilgrastim, Colony stimulating factor (e.g., Neulasta®), ribavirin and peg Intron A (e.g., Rebetrone®), Pegvisomant, PEGylated human growth hormone receptor antagonist, (e.g., Somavert®), recombinant activated protein C (e.g., Xigris®), Omalizumab, Immunoglobulin E (IgE) blocker (e.g., Xolair®) and Ibritumomab tiuxetan (murine MAb) (e.g., Zevalin®).

In various embodiments of the transgenic vertebrate animal of the present invention, the expression of the transgene may be restricted to specific subsets of cells, tissues or developmental stages utilizing, for example, trans-acting factors acting on the transcriptional regulatory region operably linked to the polypeptide-encoding region of interest of the present invention and which control gene expression in the desired pattern. Tissue-specific regulatory sequences and conditional regulatory sequences can be used to control expression of the transgene in certain spatial patterns. Moreover, temporal patterns of expression can be provided by, for example, conditional recombination systems or prokaryotic transcriptional regulatory sequences.

Another aspect of the present invention provides a method for the production of a heterologous protein capable of forming an antibody suitable for selectively binding an antigen. This method comprises a step of producing a transgenic vertebrate animal incorporating at least one transgene, the transgene encoding at least one heterologous polypeptide selected from an immunoglobulin

heavy chain variable region, an immunoglobulin heavy chain comprising a variable region and a constant region, an immunoglobulin light chain variable region, an immunoglobulin light chain comprising a variable region and a constant region, and a single-chain antibody comprising two peptide-linked immunoglobulin
5 variable regions.

In one embodiment of this method, the isolated heterologous protein is an antibody capable of selectively binding to an antigen and which may be generated by combining at least one immunoglobulin heavy chain variable region and at least one immunoglobulin light chain variable region, for example, cross-linked by at
10 least one disulfide bridge. The combination of the two variable regions generates a binding site that binds an antigen using methods for antibody reconstitution that are well known in the art.

The present invention also encompasses immunoglobulin heavy and light chains, or variants or derivatives thereof, to be expressed in separate transgenic
15 avians, and thereafter isolated from separate media including serum or eggs, each isolate comprising one or more distinct species of immunoglobulin polypeptide. The method may further comprise the step of combining a plurality of isolated heterologous immunoglobulin polypeptides, thereby producing an antibody capable of selectively binding to an antigen. In this embodiment, for instance, two or more
20 individual transgenic avians may be generated wherein one transgenic produces serum or eggs having an immunoglobulin heavy chain variable region, or a polypeptide comprising such, expressed therein. A second transgenic animal, having a second transgene, produces serum or eggs having an immunoglobulin light chain variable region, or a polypeptide comprising such, expressed therein.
25 The polypeptides from two or more transgenic animals may be isolated from their respective sera and eggs and combined in vitro to generate a binding site capable of binding an antigen.

One aspect of the present invention, therefore, concerns transgenic vertebrate animals such as transgenic birds, for example, transgenic chickens,
30 comprising a recombinant nucleic acid molecule and which may (though optionally) expresses a heterologous gene in one or more cells in the animal. Suitable methods for the generation of transgenic animals are known in the art and

are described in, for example, WO 99/19472, published April 22, 1999; WO 00/11151, published March 2, 2000; and WO 00/56932, published September 28, 2000, the disclosures of which are incorporated herein by reference in their entirety.

5 Embodiments of the methods for the production of a heterologous polypeptide by avian tissue such as oviduct tissue and the production of eggs which contain heterologous protein involve providing a suitable vector and introducing the vector into embryonic blastodermal cells together with an integrase, for example, a serine recombinase such as phiC31 integrase, so that the vector can
10 integrate into the avian genome. A subsequent step involves deriving a mature transgenic avian from the transgenic blastodermal cells produced in the previous steps. Deriving a mature transgenic avian from the blastodermal cells optionally involves transferring the transgenic blastodermal cells to an embryo and allowing that embryo to develop fully, so that the cells become incorporated into the bird as
15 the embryo is allowed to develop.

 Another alternative may be to transfer a transfected nucleus to an enucleated recipient cell which may then develop into a zygote and ultimately an adult bird. The resulting chick is then grown to maturity.

 In another embodiment, the cells of a blastodermal embryo are transfected
20 or transduced with the vector and integrase directly within the embryo. It is contemplated, for example, that the recombinant nucleic acid molecules of the present invention may be introduced into a blastodermal embryo by direct microinjection of the DNA into a stage X or earlier embryo that has been removed from the oviduct. The egg is then returned to the bird for egg white deposition,
25 shell development and laying. The resulting embryo is allowed to develop and hatch, and the chick allowed to mature.

 In one embodiment, a transgenic bird of the present invention is produced by introducing into embryonic cells such as, for instance, isolated avian blastodermal cells, a nucleic acid construct comprising an attB recombination site
30 capable of recombining with a pseudo-attP recombination site found within the nuclear genome of the organism from which the cell was derived, and a nucleic acid fragment of interest, in a manner such that the nucleic acid fragment of

interest is stably integrated into the nuclear genome of germ line cells of a mature bird and is inherited in normal Mendelian fashion. It is also within the scope of the invention that the targeted cells for receiving the transgene have been engineered to have a heterologous attP recombination site, or other recombination site, integrated
5 into the nuclear genome of the cells, thereby increasing the efficiency of recognition and recombination with a heterologous attB site.

In either case, the transgenic bird produced from the transgenic blastodermal cells is known as a "founder". Some founders can be chimeric or mosaic birds if, for example, microinjection does not deliver nucleic acid
10 molecules to all of the blastodermal cells of an embryo. Some founders will carry the transgene in the tubular gland cells in the magnum of their oviducts and will express the heterologous protein encoded by the transgene in their oviducts. If the heterologous protein contains the appropriate signal sequences, it will be secreted into the lumen of the oviduct and onto the yolk of an egg.

15 Some founders are germ-line founders. A germ-line founder is a founder that carries the transgene in genetic material of its germ-line tissue, and may also carry the transgene in oviduct magnum tubular gland cells that express the heterologous protein. Therefore, in accordance with the invention, the transgenic bird will have tubular gland cells expressing the heterologous protein and the
20 offspring of the transgenic bird will also have oviduct magnum tubular gland cells that express the selected heterologous protein. (Alternatively, the offspring express a phenotype determined by expression of the exogenous gene in a specific tissue of the avian.)

The stably modified oviduct cells will express the heterologous
25 polynucleotide and deposit the resulting polypeptide into the egg white of a laid egg. For this purpose, the expression vector will further comprise an oviduct-specific promoter such as ovalbumin or ovomucoid operably linked to the desired heterologous polynucleotide.

This description uses gene nomenclature accepted by the Cucurbit Genetics
30 Cooperative as it appears in the Cucurbit Genetics Cooperative Report 18:85 (1995), which are incorporated herein by reference in its entirety.

The disclosures of publications, patents, and published patent specifications referenced in this application are hereby incorporated by reference into the present disclosure to more fully describe the state of the art to which this invention pertains.

5 It will be apparent to those skilled in the art that various modifications, combinations, additions, deletions and variations can be made in the present invention without departing from the scope or spirit of the invention. For instance, features illustrated or described as part of one embodiment can be used in another embodiment to yield a still further embodiment. It is intended that the present
10 invention covers such modifications, combinations, additions, deletions and variations as come within the scope of the appended claims and their equivalents.

 The present invention is further illustrated by the following examples, which are provided by way of illustration and should not be construed as limiting. The contents of all references, published patents and patents cited throughout the
15 present application are hereby incorporated by reference in their entireties.

Example 1: Phage phiC31 Integrase Functions in Avian Cells.

(a) A luciferase vector bearing either an attB (SEQ ID NO: 2 shown in Fig. 10) or attP (SEQ ID NO: 3 shown in Fig. 11) site was co-transfected with an integrase
20 expression vector CMV-C31int (SEQ ID NO: 1) into DF-1 cells, a chicken fibroblast cell line. The cells were passaged several times and the luciferase levels were assayed at each passage.

 Cells were passaged every 3-4 days and one third of the cells were harvested and assayed for luciferase. The expression of luciferase was plotted as a
25 percentage of the expression measured 4 days after transfection. A luciferase expression vector bearing an attP site as a control was also included.

 As can be seen in Fig. 2, in the absence of integrase, luciferase expression from a vector bearing attP or attB decreased to very low levels after several days. However, luciferase levels were persistent when the luciferase vector bearing attB
30 was co-transfected with the integrase expression vector, indicating that the luciferase vector had stably integrated into the avian genome.

(b) A drug-resistance colony formation assay was used to quantitate integration efficiency. The puromycin resistance expression vector pCMV-pur was outfitted with an attB (SEQ ID NO: 4 shown in Fig. 12) or an attP (SEQ ID NO: 5 shown in Fig. 13) sites. Puromycin resistance vectors bearing attB sites were cotransfected
5 with phiC31 integrase or a control vector into DF-1 cells. One day after transfection, puromycin was added. Puromycin resistant colonies were counted 12 days post-transfection.

In the absence of co-transfected integrase expression, few DF-1 cell colonies were observed after survival selection. When integrase was co-expressed,
10 multiple DF-1 cell colonies were observed, as shown in Fig. 3. Similar to the luciferase expression experiment, the attB sequence (but not the attP sequence) was able to facilitate integration of the plasmid into the genome. Fig. 3 also shows that phiC31 integrase functions at both 37° Celsius and 41° Celsius. Integrase also functions in quail cells using the puromycin resistance assay, as shown in Fig. 4.

(c) The CMV-pur-attB vector (SEQ ID NO: 4) was also cotransfected with an enhanced green fluorescent protein (EGFP) expression vector bearing an attB site (SEQ ID NO: 6 shown in Fig. 14) into DF-1 cells and the phiC31 integrase expression vector CMV-C31int (SEQ ID NO: 1). After puromycin selection for 12 days, the colonies were viewed with UV light to determine the percentage of cells
15 that expressed EGFP. Approximately 20% of puromycin resistant colonies expressed EGFP in all of the cells of the colony, as shown in Fig. 5, indicating that the integrase can mediate multiple integrations per cell.

(d) PhiC31 integrase promoted the integration of large transgenes into avian cells. A puromycin expression cassette comprising a CMV promoter, puromycin resistance gene, polyadenylation sequence and the attB sequence was inserted into
25 a vector containing a 12.0 kb lysozyme promoter and the human interferon α 2b gene (SEQ ID NO: 7 shown in Fig. 15) and into a vector containing a 10.0 kb ovomucoid promoter and the human interferon α 2b gene (SEQ ID NO: 8) as shown in Fig. 16.

30 DF-1 cells were transfected with donor plasmids of varying lengths bearing a puromycin resistance gene and an attB sequence in the absence or presence of an integrase expression plasmid. Puromycin was added to the culture media to kill

those cells which did not contain a stably integrated copy of the puromycin resistance gene. Cells with an integrated gene formed colonies in the presence of puromycin in 7-12 days. The colonies were visualized by staining with methylene blue and the entire 60 mm culture dish was imaged.

- 5 PhiC31 integrase mediated the efficient integration of both vectors as shown in Fig. 7.

Example 2: Cell Culture Methods

DF-1 cells were cultured in DMEM with high glucose, 10% fetal bovine
10 serum, 2 mM L-glutamine, 100 units/ml penicillin and 100 µg/ml streptomycin at 37° Celsius and 5% CO₂. A separate population of DF-1 cells was grown at 41° Celsius. These cells were adapted to the higher temperature for one week before they were used for experiments.

Quail QT6 cells were cultured in F10 medium (Gibco) with 5% newborn
15 calf serum, 1% chicken serum heat inactivated (at 55° Celsius for 45 mins), 10 units/ ml penicillin and 10 µg/ml streptomycin at 37° Celsius and 5% CO₂.

Example 3: Selection and Assay Methods

(a) Puromycin selection assay: About 0.8×10^6 DF-1 (chicken) or QT6 (quail) cells
20 were plated in 60 mm dishes. The next day, the cells were transfected as follows:

10 to 50 ng of a donor plasmid and 1 to 10 µg of an Integrase-expressing plasmid DNA were mixed with 150 µl of OptiMEM. 15 µl of DMRIE-C was mixed with 150 µl of OptiMEM in a separate tube, and the mixtures combined and incubated for 15 mins. at room temperature.

25 While the liposome/DNA complexes were forming, the cells were washed with OptiMEM and 2.5 ml of OptiMEM was added. After 15 minutes, 300 µl of the DNA-lipid mixture was added drop wise to the 2.5 ml of OptiMEM covering the cell layers. The cells were incubated for 4-5 hours at either 37° Celsius or 41° Celsius, 5% CO₂. The transfection mix was replaced with 3 mls of culture media.
30 The next day, puromycin was added to the media at a final concentration of 1 µg/ml, and the media replaced every 2 to 4 days. Puromycin resistant colonies were counted or imaged 10-12 days after the addition of puromycin.

(b) Luciferase assay: Chicken DF-1 or quail QT6 cells (0.8×10^6) were plated in 60 mm dishes. Cells were transfected as described above. The cells from a plate were transferred to a new 100 mm plate when the plate became confluent, typically on day 3-4, and re-passaged every 3-4 days.

5 At each time point, one-third of the cells from a plate were replated, and one-third were harvested for the luciferase assay. The cells were pelleted in an eppendorf tube and frozen at -70°C .

10 The cell pellet was lysed in 200 μl of lysis buffer (25 mM Tris-acetate, pH7.8, 2mM EDTA, 0.5% Triton X-100, 5% glycerol). Sample (5 μl) was assayed using the Promega BrightGlo reagent system.

(c) Visualization of EGFP: EGFP expression was visualized with an inverted microscope with FITC illumination [Olympus IX70, 100 W mercury lamp, HQ-FITC Band Pass Emission filter cube, exciter 480/40 nm, emission 535/50 nm, 20X phase contrast objective (total magnification was $2.5 \times 10 \times 20$)].

15 (d) Staining of cell colonies: After colonies had formed, typically after 7-12 days of culture in puromycin medium, the cells were fixed in 2% formaldehyde, 0.2% glutaraldehyde for 15 mins, and stained in 0.2% methylene blue for 30 mins. followed by several washes with water. The plates were imaged using a standard CCD camera in visible light.

20

Example 4: Production of Genetically Transformed Avian Cells

Avian stage X blastodermal cells are used as the cellular vector for the transgenes. Stage X embryos are collected and the cells dispersed and mixed with plasmid DNA. The transgenes are then introduced to blastodermal cells via
25 electroporation. The cells are immediately injected back into recipient embryos.

The cells are not cultured for any time period to ensure that they remain capable of contributing to the germline of resulting chimeric embryos. However, because there is no culture step, cells that bear the transgene cannot be identified. Typically, only a small percentage of cells introduced to an embryo will bear a
30 stably integrated transgene (0.01 to 1%). To increase the percentage of cells bearing a transgene, therefore, the transgene vector bears an attB site and is co-electroporated with a vector bearing the CMV promoter driving expression of the

phiC31 transgene (CMV-C31int (SEQ ID NO: 1)). The integrase then drives integration of the transgene vector into the nuclear genome of the avian cell and increases the percentage of cells bearing a stable transgene.

(a) Preparation of avian stage X blastodermal cells:

- 5 i) Collect fertilized eggs from Barred Rock or White leghorn chickens (Gallus gallus) or quail (Japonica coturnix) within 48 hrs. of laying;
- ii) Use 70% ethanol to clean the shells;
- iii) Crack the shells and open the eggs;
- iv) Remove egg whites by transferring yolks to opposite halves of shells,
10 repeating to remove most of the egg whites;
- v) Put egg yolks with embryo discs facing up into a 10cm petri dish;
- vi) Use an absorbent tissue to gently remove egg white from the embryo
discs;
- vii) Place a Whatman filter paper 1 ring over the embryos;
- 15 viii) Use scissors to cut the membranes along the outside edge of the paper
ring while gently lifting the ring/embryos with a pair of tweezers;
- ix) Insert the paper ring with the embryos at a 45 degrees angle into a petri
dish containing PBS-G solution at room temperature;
- x) After ten embryo discs are collected, gently wash the yolks from the
20 blastoderm discs using a Pasteur pipette under a stereo microscope;
- xi) Cut the discs by a hair ring cutter (a short piece of human hair is bent
into a small loop and fastened to the narrow end of a Pasteur pipette
with Parafilm);
- xii) Transfer the discs to a 15 ml sterile centrifuge tube on ice;
- 25 xiii) Place 10 to 15 embryos per tube and allow to settle to the bottom (about
5 mins.);
- xiv) Aspirate the supernatant from the tube;
- xv) Add 5 mls of ice-cold PBS without Ca^{++} and Mg^{++} , and gently pipette 4
to 5 times using a 5 mls pipette;
- 30 xvi) Incubate in ice for 5-7 mins. to allow the blastoderms to settle, and
aspirate the supernatant;

- xvii) Add 3 mls of ice cold 0.05% trypsin/0.02% EDTA to each tube and gently pipette 3 to 5 times using a 5 ml pipette;
 - xviii) Put the tube in ice for 5 mins. and then flick the tube by finger 40 times. Repeat;
 - 5 xix) Add 0.5 mls FBS and 3-5 mls BDC medium to each tube and gently pipette 5-7 times using a 5 ml pipette;
 - xx) Spin at 500 rpm (RCF 57 x g) at 4° Celsius for 5 mins;
 - xxi) Remove the supernatant and add 2 mls ice cold BDC medium into each tube; and
 - 10 xxii) Resuspend the cells by gently pipetting 20-25 times; and
 - xxiii) Determine the cell titer by hemacytometer and ensure that about 95% of all BDCs are single cells, and not clumped.
- (b) Transfection of linearized plasmids into blastodermal cells by small scale electroporation:
- 15 i) Centrifuge the blastodermal cell suspension from step (xxiii) above at RCF 57 x g, 4° Celsius, for 5 mins;
 - ii) Resuspend cells to a density of $1-3 \times 10^6$ per ml with PBS without Ca^{2+} and Mg^{2+} ;
 - iii) Add linearized DNA, 1-30 μg per $1-3 \times 10^5$ blastodermal cells in an eppendorf tube at room temperature. Add equimolar molar amounts of the non-linearized transgene plasmid bearing an attB site, and an integrase expression plasmid;
 - 20 iv) Incubate at room temperature for 10 mins;
 - v) Aliquot 100 μl of the DNA-cell mixture to a 0.1 cm cuvette at room temperature;
 - 25 vi) Electroporate at 240 V and 25 μFD (or 100 V and 125 μFD for quail cells) using, for example, a Gene Pulser II™ (BIO-RAD).
 - vii) Incubate the cuvette at room temperature for 1-10 mins.
 - viii) Before the electroporated cells are injected into a recipient embryo, they are transferred to a eppendorf tube at room temperature. The cuvette is washed with 350 μl of media, which is transferred to the eppendorf, spun at room temperature and re-suspended in 0.01-0.3 ml medium;
 - 30

- ix) Inject 1-10 μ l of cell suspension into the subgerminal cavity of a non-irradiated or, for example, an irradiated (e.g., with 300-900 rads) stage X egg. Shell and shell membrane are removed and, after injection, resealed according to U.S. Patent No. 6,397,777, issued June 6, 2002, the disclosure of which is incorporated herein by reference in its entirety; and
 - x) The egg is then incubated to hatching.
- (c) Blastodermal Cell Culture Medium:
- i) 409.5 mls DMEM with high glucose, L-glutamine, sodium pyruvate, pyridoxine hydrochloride;
 - ii) 5 mls Men non-essential amino acids solution, 10 mM;
 - iii) 5 mls Penicillin-streptomycin 5000 U/ml each;
 - iv) 5 mls L-glutamine, 200 mM;
 - v) 75 mls fetal bovine serum; and
 - vi) 0.5 mls β -mercaptoethanol, 11.2mM.

Example 5: Transfection of Stage X Embryos with attB Plasmids

(a) DNA-PEI: Twenty-five μ g of a phage phiC31 integrase expression plasmid (pCMV-int), and 25 μ g of a luciferase-expressing plasmid (p β -actin-GFP-attB) are combined in 200 μ l of 28 mM Hepes (pH 7.4). The DNA/Hepes is mixed with an equal volume of PEI which has been diluted 10-fold with water. The DNA/Hepes/PEI is incubated at room temperature for 15 mins. Three to seven μ l of the complex are injected into the subgerminal cavity of windowed stage X white leghorn eggs which are then sealed and incubated as described in U.S. Patents No. 6,397,777, issued June 6, 2002. The complexes will also be incubated with blastodermal cells isolated from stage X embryos which are subsequently injected into the subgerminal cavity of windowed irradiated stage X white leghorn eggs. Injected eggs are sealed and incubated as described above.

(b) Adenovirus-PEI:

Two μ g of a phage phiC31 integrase expression plasmid (pCMV-int), 2 μ g of a GFP expressing plasmid (p β -actin-GFP-attB) and 2 μ g of a luciferase expressing plasmid (pGLB) were incubated with 1.2 μ l of JetPEITM in 50 μ l of 20 mM Hepes

buffer (pH7.4). After 10 mins at 25°C, 3×10^9 adenovirus particles (Ad5-Null, Qbiogene) were added and the incubation continued for an additional 10 mins. Embryos are transfected in ovo or ex ovo as described above.

5

Example 6: Stage I Cytoplasmic Injection

Production of transgenic chickens by cytoplasmic DNA injection using DNA injection directly into the germinal disk as described in Sang et al, Mol. Reprod. Dev., 1: 98-106 (1989); Love et al, Biotechnology, 12: 60-63 (1994) incorporated herein by reference in their entirety.

10

In the method of the present invention, fertilized ova, or stage I embryos, are isolated from euthanized hens 45 mins. to 4 hrs. after oviposition of the previous egg. Alternatively, eggs were isolated from hens whose oviducts have been fistulated according to the techniques of Gilbert & Wood-Gush, J. Reprod. Fertil., 5: 451-453 (1963) and Pancer et al, Br. Poult. Sci., 30: 953-7 (1989)

15

incorporated herein in their entirety.

An isolated ovum was placed in dish with the germinal disk upwards. Ringer's buffer medium was then added to prevent drying of the ovum. Any suitable microinjection assembly and methods for microinjecting and reimplanting avian eggs are useful in the method of cytoplasmic injection of the present invention. A particularly suitable apparatus and method for use in the present invention is described in U.S. Patent Application Serial No: 09/919,143, published July 31, 2001, the disclosure of which is incorporated in its entirety herein by reference. The avian microinjection system described in the '143 Application allowed the loading of a DNA solution into a micropipette, followed by prompt positioning of the germinal disk under the microscope and guided injection of the DNA solution into the germinal disk. Injected embryos could then be surgically transferred to a recipient hen as described, for example, in Olsen & Neher, J. Exp. Zool., 109: 355-66 (1948) and Tanaka et al, J. Reprod. Fertil., 100: 447-449 (1994). The embryo was allowed to proceed through the natural in vivo cycle of albumin deposition and hard-shell formation. The transgenic embryo is then laid as a hard-shell egg which was incubated until hatching of the chick. Injected embryos were surgically transferred to recipient hens via the ovum transfer method

of Christmann et al in PCT/US01/26723, published August 27, 2001, the disclosure of which is incorporated herein by reference in its entirety, and hard shell eggs were incubated and hatched.

5 Approximately 25 nl of DNA solution (about 60ng/ μ l) with either integrase mRNA or protein were injected into a germinal disc of stage I White Leghorn embryos obtained 90 minutes after oviposition of the preceding egg. Typically the concentration of integrase mRNA used was 100 ng/ μ l, and the concentration of integrase protein was 66 ng/ μ l.

10 To synthesize the integrase mRNA, a plasmid template encoding the integrase protein was linearized at the 3' end of the transcription unit. mRNA was synthesized, capped and a polyadenine tract added using the mMACHINE mMACHINE T7 Ultra Kit™ (Ambion, Austin, TX). The mRNA was purified by extraction with phenol and chloroform and precipitated with isopropanol. The integrase protein was expressed in E. coli and purified as described by Thorpe et al,
15 Mol. Microbiol., 38: 232-241 (2000).

A plasmid encoding for the integrase protein is transfected into the target cells. However, since the early avian embryo transcriptionally silent until it reaches about 22,000 cells, injection of the integrase mRNA or protein was expected to result in better rates of transgenesis, as shown in the Table 1 below.

20 The chicks produced by this procedure were screened for the presence of the injected transgene using a high throughput PCR-based screening procedure as described in Harvey et al, Nature Biotech., 20: 396-399 (2002).

Table 1: Summary of cytoplasmic injection results using different integrase strategies

Experimental group	Ovum transfers	Hard shells produced (%)	Chicks hatched (%) *	Transgenic chicks (%) ‡
No Integrase	5164	3634 (70%)	500 (14%)	58 (11.6%)
Integrase mRNA	1109	833 (75%)	115 (13.8%)	19 (16.5%)
Integrase protein	374	264 (70.6%)	47(17.8%)	16 (34%)

* : Percentages based on the number of hard shells

‡ : Percentages based on the number of hatched birds

5

Example 7: Characterization of phiC31 Integrase-Mediated Integration Sites in the Chicken Genome

To characterize phiC31-mediated integration into the chicken genome, a plasmid rescue method was used to isolate integrated plasmids from transfected and selected chicken fibroblasts. Plasmid pCR-XL-TOPO-CMV-pur-attB (SEQ ID NO: 10, shown in Fig. 18) does not have BamH I or Bgl II restriction sites. Genomic DNA from cells transformed with pCR-XL-TOPO-CMV-pur-attB was cut with BamH I or Bgl II (either or both of which would cut in the flanking genomic regions) and religated so that the genomic DNA surrounding the integrated plasmid would be captured into the circularized plasmid. The flanking DNA of a number of plasmids were then sequenced.

DF-1 cells (chicken fibroblasts), 4×10^5 were transfected with 50 ng of pCR-XL-TOPO-CMV-pur-attB and 1 μ g of pCMV-int. The following day, the culture medium was replaced with fresh media supplemented with 1 μ g/ml puromycin. After 10 days of selection, several hundred puromycin-resistant colonies were evident. These were harvested by trypsinization, pooled, replated on 10 cm plates and grown to confluence. DNA was then extracted.

Isolated DNA was digested with BamH I and Bgl II for 2-3 hrs, extracted with phenol:chloroform:isoamyl alcohol chloroform:isoamyl alcohol and ethanol precipitated. T4 DNA ligase was added and the reaction incubated for 1 hr at room

temperature, extracted with phenol:chloroform:isoamyl alcohol and chloroform:isoamyl alcohol, and precipitated with ethanol. 5 µl of the DNA suspended in 10µl of water was electroporated into 25 µl of GenehogsTM (Invitrogen) in an 0.1 cm cuvette using a GenePulser II (Biorad) set at 1.6 kV, 100
5 ohms, 25 uF and plated on Luria Broth (LB) plates with 5 µg/ml phleomycin (or 25 µg/ml zeocin) and 20 µg/ml kanamycin. Approximately 100 individual colonies were cultured, the plasmids extracted by standard miniprep techniques and digested with Xba I to identify clones with unique restriction fragments.

Thirty two plasmids were sequenced with the primer attB-for (5'-
10 TACCGTCGACGATGTAGGTCACGGTC-3') (SEQ ID NO: 12) which allows sequencing across the crossover site of attB and into the flanking genomic sequence. All of plasmids sequenced had novel sequences inserted into the crossover site of attB, indicating that the clones were derived from plasmid that had integrated into the chicken genome via phiC31 integrase-mediated
15 recombination.

The sequences were compared with sequences at GenBank using Basic Local Alignment Search Tool (BLAST). Most of the clones harbored sequences homologous to Gallus genomic sequences in the TRACE database.

20 **Example 8: Insertion of a Wild-Type attP Site into the Avian Genome**
Augments Integrase-Mediated Integration and Transgenesis

The chicken B-cell line DT40 cells (Buerstedde et al, E.M.B.O. J., 9: 921-927 (1990)) are useful for studying DNA integration and recombination processes (Buerstedde & Takeda, Cell, 67:179-88 (1991)). DT40 cells were engineered to
25 harbor a wild-type attP site isolated from the Streptomyces phage phiC31. Two independent cell lines were created by transfection of a linearized plasmid bearing an attP site linked to a CMV promoter driving the resistance gene to G418 (DT40-NLB-attP) or bearing an attP site linked to a CMV promoter driving the resistance gene for puromycin (DT40-pur-attP). The transfected cells were cultured in the
30 presence of G418 or puromycin to enrich for cells bearing an attP sequence stably integrated into the genome.

A super-coiled luciferase vector bearing an attB (SEQ ID NO: 2 shown in Fig. 10) was co-transfected, together with an integrase expression vector CMV-C31int (SEQ ID NO: 1) or a control, non-integrase expressing vector (CMV-BL) into wild-type DT40 cells and the stably transformed lines DT40-NLB-attP and DT40-pur-attP.

Cells were passaged at 5, 7 and 14 days post-transfection and about one third of the cells were harvested and assayed for luciferase. The expression of luciferase was plotted as a percentage of the expression measured 5 days after transfection. As can be seen in Fig. 21, in the absence of integrase, or in the presence of integrase but in the DT40 cells lacking an inserted wild-type attP site, luciferase expression from a vector bearing attB progressively decreased to very low levels. However, luciferase levels were persistent when the luciferase vector bearing attB was co-transfected with the integrase expression vector into the attP bearing cell lines DT40-NLB-attP and DT40-pur-attP. Inclusion of an attP sequence in the avian genome augments the level of integration efficiency beyond that afforded by the utilization of endogenous pseudo-attP sites.

Example 9: Generation of attP Transgenic Cell Line and Birds Using an NLB Vector

The NLB-attP retroviral vector is injected into stage X chicken embryos laid by pathogen-free hens. A small hole is drilled into the egg shell of a freshly laid egg, the shell membrane is cut away and the embryo visualized by eye. With a drawn needle attached to a syringe, 1 to 10 μ l of concentrated retrovirus, approximately 2.5×10^5 IU, is injected into the subgerminal cavity of the embryo. The egg shell is resealed with a hot glue gun. Suitable methods for the manipulation of avian eggs, including opening and resealing hard shell eggs are described in U.S. Patent Serial Nos: 5,897,998, issued May 27, 1999 and 6,397,777, issued June 4, 2002, the disclosures of which are herein incorporated by reference in their entireties.

Typically, 25% of embryos hatch 21 days later. The chicks are raised to sexual maturity and semen samples are taken. Birds that have a significant level of the transgene in sperm DNA will be identified, typically by a PCR-based assay.

Ten to 25% of the hatched roosters will be able to give rise to G1 transgenic offspring, 1 to 20% of which may be transgenic. DNA extracted from the blood of G1 offspring is analyzed by PCR and Southern analysis to confirm the presence of the intact transgene. Several lines of transgenic roosters, each with a unique site of attP integration, are then bred to non-transgenic hens, giving 50% of G2 transgenic offspring. Transgenic G2 hens and roosters from the same line can be bred to produce G3 offspring homozygous for the transgene. Homozygous offspring will be distinguished from hemizygous offspring by quantitative PCR. The same procedure can be used to integrate an attB or attP site into transgenic birds.

10

Example 10: Expression of Immunoglobulin Chain Polypeptides by Transgenic Chickens

Bacterial artificial chromosomes (BACs) containing a 70 kb segment of the chicken ovomucoid gene with the light and heavy chain cDNAs for a human monoclonal antibody inserted along with an internal ribosome entry site into the 3' untranslated region of the ovomucoid gene were equipped with the attB sequence. The heavy and light chain cDNAs were inserted into separate ovomucoid BACs such that expression of an intact monoclonal antibody requires the presence of both BACs in the nucleus.

Several hens produced by coinjection of the attB-bearing ovomucoid BACs and integrase-encoding mRNA into stage I embryos produced intact monoclonal antibodies in their egg white. One hen, which had a high level of the light chain ovomucoid BAC in her blood DNA as determined by quantitative PCR particularly expressed the light chain portion of the monoclonal antibody in the egg white at a concentration of 350 nanograms per ml, or approximately 12 µg per egg.

25

Example 11: Stage I Cytoplasmic Injection with Integrase Activity and PEI

Production of transgenic chickens by cytoplasmic DNA injection directly into the germinal disk was done as described in Example 6.

DNA (about 60ng/µl) which includes a transgene was placed in approximately 25 nl of aqueous solution with integrase mRNA or integrase protein and was mixed with an equal volume of PEI that had been diluted ten fold. The

30

mixture was injected into a germinal disc of stage I White Leghorn embryos obtained about 90 minutes after oviposition of the preceding egg. Typically the concentration of integrase mRNA used was about 100 ng/ μ l, and the concentration of integrase protein was about 66 ng/ μ l. The integrase mRNA was synthesized according to Example 6.

Transgenic chicks produced by this procedure using: integrase mRNA/PEI and integrase protein/PEI showed positive results for the presence of heterologously expressed protein in the blood, semen and egg white.

Example 12: Stage I Cytoplasmic Injection with Integrase Activity and NLS

Production of transgenic chickens by cytoplasmic DNA injection directly into the germinal disk was done as described in Example 6.

DNA which includes a transgene was suspended in 0.25 M KCl and SV40 T antigen nuclear localization signal peptide (NLS peptide, amino acid sequence CGGPKKKRKVG (SEQ ID NO: 13)) was added to achieve a peptide DNA molar ratio of 100:1. The DNA (about 60ng/ μ l) was allowed to associate with the SV40 T antigen NLS peptide by incubating at 25 degrees C for about 15 minutes.

Integrase mRNA or integrase protein was added to approximately 25 nl of an aqueous DNA/NLS solution, typically, to produce a final concentration of integrase mRNA of about 50 ng/ μ l, or an integrase protein concentration of about 33 ng/ μ l. The mixture was injected into a germinal disc of stage I White Leghorn embryos obtained about 90 minutes after oviposition of the preceding egg. The integrase mRNA was synthesized as according to Example 6.

Transgenic chicks produced by this procedure using: integrase mRNA/NLS and integrase protein/NLS showed positive results for the presence of heterologously expressed protein in blood, semen and egg white.

Example 13: Dispersing of Plasmid DNA in Avian Stage I Embryos

DNA samples are Cy3 labeled with a Cy3 ULS labeling kit (Amersham Pharmacia Biotech). Briefly, plasmid DNA (1 μ g) was sheared to approximately 100 to 500 bp fragments by sonication. Resulting DNA was incubated at 65°C for 15 min in Cy3 ULS labeling solution and unincorporated Cy3 dye was removed by

spin column chromatography (CentriSep, Princeton Separations). The distribution of the DNA in stage I avian embryos was visualized after introduction into the stage I avian embryo. Enough high molecular weight or low molecular weight PEI was added to the DNA to coat the DNA. Typically, PEI was added to the DNA to a concentration of about 5%. Any useful volume of DNA/PEI can be used, for example about 25 nl.

Figure 22 shows an avian stage one embryo containing Cy3 labeled naked DNA. In Figure 22 it can be seen that the DNA is localized to certain areas of the embryo. Figure 23 and Figure 24 show an avian stage one embryo containing Cy3 labeled DNA coated with low molecular (22 kD) weight PEI (Figure 23) and high molecular weight (25 kD) PEI (Figure 24). In Figures 23 and 24, it can be seen that the DNA is dispersed throughout the embryos.

These experiments show that DNA/PEI conjugates are distributed more uniformly in the cytoplasm of injected embryos when compared with naked DNA.

Example 14: Production of an attP Transgenic Chicken

G0 transgenic chickens have been produced as described in Example 9. Several hundred stage X White Leghorn eggs were injected with the NLB-attP vector and about 50 chicks hatched. Sperm from approximately 30% of the hatched roosters has been shown to be positive for the attP site. These hemizygotic chickens are used to generate transgenic G2 chickens homozygotic for the attP site.

Example 15: Cytoplasmic Injection of attP Stage I Embryos with OMC24-attB-IRES-CTLA4

Transgenic chickens are produced by cytoplasmic DNA injection directly into the germinal disk of eggs laid by transgenic homozygous attP chickens and fertilized with sperm from the same line of homozygous attP roosters, the line produced as described in Example 14. The cytoplasmic injections are carried out as described in U.S. Patent Application Serial No. 09/919,143, filed July 31, 2001, ('143 Application) and U.S. Patent Application Serial No. 10/251,364, filed September 18, 2002. The disclosures of each of these two patent applications are incorporated herein by reference in their entirety.

Stage I embryos are isolated 45 mins. to 4 hrs. after oviposition of the previous egg. An isolated embryo is placed in a dish with the germinal disk upwards. Ringer's buffer medium is added to prevent drying of the ovum. The avian microinjection system described in the '143 Application allows for the loading of DNA solution into a micropipette, followed by prompt positioning of the germinal disk under the microscope and guided injection of the DNA solution into the germinal disk.

Approximately 25 nl of a DNA solution (about 60ng/ μ l) of the 77 kb OMC24-attB-IRES-CTLA4, disclosed in US Patent Application No. 10/856,218, filed May 28, 2004, the disclosure of which is incorporated in its entirety herein by reference, with either integrase mRNA or protein are injected into a germinal disc of the isolated stage I embryos. Typically, the concentration of integrase mRNA used is 100 ng/ μ l or the concentration of integrase protein is 66 ng/ μ l.

To synthesize the integrase mRNA, a plasmid template encoding the integrase protein is linearized at the 3' end of the transcription unit. mRNA is synthesized, capped and a polyadenine tract added using the mMACHINE mMESSAGE mMACHINE T7 Ultra Kit™ (Ambion, Austin, TX). The mRNA is purified by extraction with phenol and chloroform and precipitated with isopropanol. The integrase protein is expressed in E. coli and purified as described by Thorpe et al, Mol. Microbiol., 38: 232-241 (2000).

Injected embryos are surgically transferred to a recipient hen as described in Olsen & Neher, J. Exp. Zool., 109: 355-66 (1948) and Tanaka et al, J. Reprod. Fertil., 100: 447-449 (1994). The embryo is allowed to proceed through the natural in vivo cycle of albumin deposition and hard-shell formation. The transgenic embryo is then laid as a hard-shell egg which is incubated until hatching of the chick. Injected embryos are surgically transferred to recipient hens via the ovum transfer method of Christmann et al in PCT/US01/26723, published August 27, 2001, the disclosure of which is incorporated by reference in its entirety, and hard shell eggs are incubated and hatched.

The chicks produced by this procedure are screened for the presence of the injected transgene using a high throughput PCR-based screening procedure as described in Harvey et al, Nature Biotech., 20: 396-399 (2002). Approximately

20% of the chicks are positive for the transgene. Eggs from each of the mature hens carrying the transgene are positive for CTLA4.

Example 16: Cytoplasmic Injection of attP Stage I Chicken Embryos with OM10-attB-CTLA4

Transgenic chickens are produced by cytoplasmic DNA injection directly into the germinal disk of eggs laid by transgenic homozygous attP chickens and fertilized with sperm from the same line of homozygous attP roosters essentially as described in Example 15.

Approximately 25 nl of a 60ng/ μ l DNA solution of the OMC24-attB-IRES-CTLA4 construct of Example 15 with the OMC24 70 kb ovomucoid gene expression controlling region and IRES of the construct replaced with the 10 kb ovomucoid gene expression controlling region of pBS-OVMUP-10, also disclosed in US Patent Application No. 10/856,218, filed May 28, 2004, is injected into a fertilized germinal disc of stage I embryos along with and integrase protein. The concentration of integrase protein used is 66 ng/ μ l.

Injected embryos are then surgically transferred to a recipient hen, hard shell eggs are produced, incubated and hatched. Approximately 30% of the chicks are positive for the transgene. Eggs from each of the mature hens carrying the transgene are positive for CTLA4.

Example 17: Production of attP Transgenic Quail Using an NLB vector

The NLB-attP retroviral vector is injected into stage X quail embryos laid by pathogen-free quail. A small hole is drilled into the egg shell of a freshly laid egg, the shell membrane cut away and the embryo visualized by eye. With a drawn needle attached to a syringe, 1 to 10 μ l of concentrated retrovirus, approximately 1.0×10^5 IU, is injected into the subgerminal cavity of the embryo. The egg shell is resealed with a hot glue gun

Typically, 25% of embryos hatch. The chicks are raised to sexual maturity and semen samples are taken. Birds that have a significant level of the transgene in their sperm DNA will be identified, typically by a PCR-based assay. Of the

hatched G0 male quail, about 1% to about 20% are transgenic. The transgenic G0 quail are bred to nontransgenic quail to produce hemizygotic G1 offspring. DNA extracted from the blood of G1 offspring is analyzed by PCR and Southern analysis to confirm the presence of the intact transgene. Several lines of hemizygotic
5 transgenic male quail, each with a unique site of attP integration, are then bred to non-transgenic quail giving G2 offspring, 50% of which are transgenic. Transgenic G2 male and female from the same line are then bred to produce G3 offspring homozygous for the transgene. Homozygous offspring are distinguished from hemizygous offspring by quantitative PCR.

10

**Example 18: Cytoplasmic Injection of attP Stage I Quail Embryos with
OMC24-attB-IRES-G-CSF**

Transgenic quail are produced by cytoplasmic DNA injection directly into the germinal disk of eggs laid by fully transgenic homozygous attP quail produced
15 as described in Example 17. The cytoplasmic injections are carried out essentially as described in the '143 Application and U.S. Patent Application Serial No. 10/251,364, filed September 18, 2002.

Stage I embryos from homozygous attP quail fertilized with sperm from a homozygous attP quail are isolated approximately 90 minutes after oviposition of
20 the previous egg. An isolated embryo is placed in a dish with the germinal disk upwards. Ringer's buffer medium is added to prevent drying of the ovum. The avian microinjection system described in the '143 Application is used to inject approximately 25 nl of a DNA solution (about 60ng/ μ l) of OMC24-attB-IRES-CTLA4, with the CTLA coding sequence replaced with the coding sequence for a
25 human-granulocyte colony stimulating factor, and integrase protein into the germinal disc of the stage I quail embryos. The concentration of integrase protein used is 66 ng/ μ l.

Injected embryos are surgically transferred to a recipient quail essentially as described in Olsen & Neher, J. Exp. Zool., 109: 355-66 (1948) and Tanaka et al, J.
30 Reprod. Fertil., 100: 447-449 (1994). The embryo is allowed to proceed through the natural in vivo cycle of albumin deposition and hard-shell formation. The

transgenic embryo is then laid as a hard-shell egg which is incubated until hatching of the chick.

The chicks produced by this procedure are screened for the presence of the injected transgene using a high throughput PCR-based screening procedure as described in Harvey et al, Nature Biotech., 20: 396-399 (2002). Approximately 20% of the chicks are positive for the transgene. Eggs from each of the mature female quail carrying the transgene are positive for G-CSF.

Example 19: Generation of attP Transgenic Duck

10

Using an NLB vector

The NLB-attP retroviral vector is injected into stage X Duck embryos laid by pathogen-free Ducks. A small hole is drilled into the egg shell of a freshly laid egg, the shell membrane cut away and the embryo visualized by eye. About 1 to 10 μ l of concentrated retrovirus, approximately 2.5×10^5 IU, is injected into the subgerminal cavity of the embryo. The egg shell is resealed with a hot glue gun

Homozygous G3 offspring are obtained essentially as described in Example 17 for quail.

Example 20: Stage I Cytoplasmic Injection of attP Stage I Duck Embryos with

20

OM24-attB-IRES-CTLA4

Transgenic ducks are produced by cytoplasmic DNA injection directly into the germinal disk of eggs laid by homozygous attP ducks fertilized with sperm from homozygous attP ducks. The injection of the stage I embryos is carried out essentially as described in the '143 Application and U.S. Patent Application Serial No. 10/251,364, filed September 18, 2002. Approximately 25 nl of a DNA solution (about 60ng/ μ l) of OMC24-attB-IRES-CTLA4, with the CTLA4 coding region replaced with a coding sequence for human erythropoietin, and integrase encoding mRNA and protein is injected into the germinal disc of the stage I embryos. The concentration of integrase mRNA used is 100 ng/ μ l. The injected embryos are surgically transferred to a recipient duck and the embryo is allowed to proceed through the natural in vivo cycle of albumin deposition and hard-shell formation. The transgenic embryo is laid as a hard-shell egg which is incubated

until hatching and the chicks are screened for the presence of the injected transgene. Approximately 20% of the chicks are positive for the transgene. Eggs from each of the mature female ducks carrying the transgene are positive for erythropoietin.

5

Example 21: Production of Transchromosomic Chickens Using Satellite

DNA-Based Artificial Chromosomes

Satellite DNA-based artificial chromosomes (ACEs, as described in Lindenbaum et al Nucleic Acids Res (2004) vol 32 no. 21 e172) were isolated by a
10 dual laser high-speed flow cytometer as described previously (de Jong, G, et al. Cytometry 35: 129-133, 1999).

The flow-sorted chromosomes were pelleted by centrifugation of a 750 μ l sample containing approximately 10^6 chromosomes at 2500 x g for 30 min at 4°C. The supernatant, except the bottom 30 microliters (μ l) containing the
15 chromosomes, was removed resulting in a concentration of about 7000 to 11,500 chromosomes per μ l of injection buffer (Monteith, et al. Methods Mol Biol 240: 227-242, 2004). Depending on the number of chromosomes to be injected, 25-100 nanoliters (nl) of injection buffer was injected per embryo.

Embryos for this study were collected from 24-36 week-old hens from
20 commercial White Leghorn variety of *G. gallus*. Embryo donor hens were inseminated weekly using pooled semen from roosters of the same breed to produce eggs for injection.

On the day of egg collection, fertile hens were euthanized 2h post oviposition by cervical dislocation. Typically, oviposition is followed by ovulation
25 of the next egg after around 24 minutes (Morris, Poultry Science 52: 423-445, 1973). The recently ovulated and fertilized eggs were collected from the upper magnum region of the oviduct under sterile conditions and placed in a glass well and covered with Ringers' Medium (Tanaka, et al. J Reprod Fertil 100: 447-449, 1994) and maintained at 41°C until microinjection.

30 Cytoplasmic injection of artificial chromosomes was achieved using the microinjection apparatus disclosed in US Patent Application No. 09/919,143, filed July 31, 2001. Chromosomes were injected into the Stage I embryos at a single

site. Each embryo was cytoplasmically injected with approximately: 175, 250, 350, 450, 550, 800 or >1000 chromosomes. The chromosomes were injected in a suspension of 25-100 nanoliters (nl) of injection buffer.

Following microinjection, the embryos were transferred to the oviduct of recipient hens using an optimized ovum transfer (OT) procedure (Olsen, M and Neher, B. J Exp Zool 109: 355-66, 1948), with the exception that the hens were anesthetized by Isoflurane gas. Typically, about 26h after OT, the recipient hens lay a hard shell egg containing the manipulated ovum. Eggs were incubated for 21 days in a regular incubator until hatching of the birds.

The chromosomes were injected into the embryos over a 9 day period. The chromosomes were divided into three batches for delivery to the embryos each batch being injected over a three day period. Chromosomes were introduced into the embryos by a single injection using the microinjection assembly disclosed in the '143 patent application. Following injection, each egg was transferred to a recipient hen. A total of 301 transfers were performed, resulting in 226 (75%) hard shells and 87 hatched chicks (38%, see Table 2).

Table 2: Hatching of embryos microinjected with satellite DNA-based artificial chromosomes.

	Ovum transfers	Hard shells produced	hatched birds
1 st batch	71	53	15
2 nd batch	113	80	33
3 rd batch	117	93	39
Totals	301	226 (75%)	87 (38%)

Previous experiments have determined that hatching is not significantly affected when embryos were injected with up to 100nl of injection buffer. Satellite DNA-based artificial chromosomes were injected in suspensions of between 25-100nl of injection buffer.

As discussed, the embryos were injected with one of seven different numbers of artificial chromosomes. There was shown to be a correlation between the number of chromosomes injected per egg and the hatch rate. All transchromosomic birds in the present study were obtained from embryos injected

with 550 chromosomes or less (see Table 3). There was no significant difference in the hatching rates observed between the experimental groups (batches 1, 2 and 3).

Six transchromosomic founders were produced based on two separate PCR analysis (6.8%, see Table 3) using primers which anneal to the puromycin resistance gene (about 75 copies of the pur^R gene are present on the chromosome. All positive birds appear normal.

Table 3: Effect of the number of Chromosomes injected per embryo on hatching and number of transchromosomic birds produced.

# chromosomes injected per embryo	# of hard shells	# chicks hatched	# of positive birds (bird tag #)
175	31	11 (35%)	3 (BB7478, BB7483, BB7515)
250	51	25 (49%)	1 (BB 7499)
350	15	6 (40%)	0
450	31	11 (35%)	0
550	39	17 (43%)	2 (BB7477, BB7523)
800	26	5 (19%) *	0
1000	33	10 (30%) *	0
Totals	226	87 (38%)	6 (6.8%)

*: hatching rates of embryos injected with >550 chromosomes was significantly lower ($p < 0.025$)

To confirm the PCR results, erythrocytes from all PCR-positive birds as well as fibroblast cells derived from skin biopsies of 5 PCR-positive birds were analyzed by interphase and metaphase FISH using a mouse-specific major satellite DNA probe (Co, et al. Chromosome Res 8: 183-191, 2000). Five of the six chicks (5.3% out of total number of chicks analyzed) tested by FISH were positive in at least one cell type (see Table 4) at 3 weeks of age. FISH analysis of erythrocytes was repeated when the birds reached 8 weeks of age and had tripled their body weight. Similar numbers of artificial chromosome-positive cells found in each bird were observed in this second FISH analysis.

Table 4: Summary of FISH analysis of Red Blood Cells (RBCs) and fibroblast cells derived from transchromosomic birds. Fibroblast cells from hen # 7515 were not available for analysis.

Bird #	Sex of Bird	% of artificial chromosome positive RBCs by FISH	% of artificial chromosome positive fibroblasts by FISH
BB7499	Female	77%	87%
BB7483	Female	0.8%	0%
BB7477	Male	3%	2.8%
BB7478	Male	15%	3%
BB7515	Female	1.3%	NA
BB7523	Male	0%	0%
Neg. control	-	0%	0%

5

To verify the chromosomes were intact, metaphase spreads from fibroblast cells derived from founders were made as described previously (Garside and Hillman (1985) *Experientia* 41: 1183-1184). FISH analysis of metaphase spreads using the major satellite DNA probe showed the artificial chromosomes appear intact, with no apparent fragmentation or translocation onto the chicken's chromosomes. FISH analysis using a mouse minor satellite probe, which detects the centromeric region of the introduced chromosomes (Wong and Rattner (1988) *J. Nucleic Acids Res* 16: 11645-11661), demonstrated the centromere of the chromosomes was intact. Furthermore, the percentage of satellite DNA-based artificial chromosomes -positive cells from metaphase spreads agreed closely to those observed in interphase FISH.

Analysis of G1 embryos from test bird BB7499 has shown the artificial chromosome to be transmitted through the germline. In addition, sperm from BB7499 was shown to test positive for the artificial chromosome which will also provide for germline transmission of the artificial chromosome.

Example 22: Production of EPO and G-CSF Vectors for the Production of Transchromosomic Chickens

25

Two vectors were constructed for introduction into Satellite DNA-based artificial chromosomes. 1OMC24-IRES1-EPO - ChromattB was constructed by inserting an EPO coding sequence into an OMC24-IRES BAC clone disclosed in

US Patent Application No. 10/856,218, filed May 28, 2004, the disclosure of which is incorporated in its entirety herein by reference. The EPO coding sequence was inserted in the clone so as to be under the control of the ovomucoid promoter. That is, the EPO coding sequence was inserted in place of the LC portion of OMC-IRES-LC. An attB site and a hygromycin^R coding sequence were also inserted into the vector in such a manner as to facilitate recombination into an attP site in a SATAC artificial chromosome (i.e., ACE), as see in FIG. 25. The attP site in the SATAC is located adjacent to an SV40 promoter which provides for expression of the hygromycin^R coding sequence upon integration of the vector into the attP site allowing for selection of cells containing a recombinant artificial chromosome (see, for example, US Patent No. 6,743,967, issued June 1, 2004; US Patent No. 6,025,155, issued February 15, 2000 and Lindenbaum et al Nucleic Acids Res (2004) vol 32 no. 21 e172 (see FIG. 25), the disclosure of each of these two patents and the publication are incorporated in their entirety herein by reference).

A coding sequence for G-CSF, which was codon optimized for expression in chicken tubular gland cells, was inserted in the 1OMC24-IRES1-EPO - ChromattB construct in place of the EPO coding sequence to produce 1OMC24-IRES-GCSF - ChromattB.

Example 23: Production of Erythropoietin and G-CSF Using Artificial Chromosomes in Chickens

Cells containing the recombinant artificial chromosome are produced and identified as described in Lindenbaum et al Nucleic Acids Res (2004) vol 32 no. 21 e172. Briefly, 2.5 µg of 1OMC24-IRES1-EPO ChromattB and 2.5 µg of an expression vector which contains a lambda integrase gene (int) having a codon mutation at position 174 to substitute a lysine for a glutamine (pCXLamROK, see Lindenbaum et al Nucleic Acids Res (2004) vol 32 no. 21 e172) are transfected by standard lipofection methodologies into LMTK- cells which contain the platform SATAC (ACE) (A of FIG. 25). Hygromycin resistant cells clones are identified by standard antibiotic selection methodologies.

Recombinant chromosomes are prepared from the cells and isolated by flow cytometry. The substantially purified artificial chromosomes are introduced into

chickens by microinjection into stage one embryos as disclosed in US Patent Application Nos. 10/679,034, filed October 2, 2003 and 09/919,143, filed July 31, 2001. Resulting chimeric germline transchromosomal avians can be identified by any useful method such as Southern blot analysis.

5

Example 24: Production of a Monoclonal Antibody Using Drosophila
Artificial Chromosomes in Turkey

Artificial chromosomes comprising a Drosophila chromosome centromere (DAC) are prepared essentially using methods described in US Patent No. 6,025,155, issued February 15, 2000, the disclosure of which is incorporated in its entirety herein by reference.

An attB site and a hygromycin^R coding sequence are inserted into the OMC24-IRES-LC and OMC24-IRES-HC vectors disclosed in US Patent Application No. 10/856,218, filed July 31, 2001, the disclosure of which is incorporated in its entirety herein by reference, which are then each cloned into a DAC essentially as described in Examples 22 and 23. The recombinant DACs are prepared and then isolated by a dual laser high-speed flow cytometer.

The flow-sorted chromosomes are pelleted by centrifugation and are diluted to a concentration of about 7000-12,000 chromosomes per μ l of injection buffer. Approximately 50 nanoliters (nl) of injection buffer is injected per turkey embryo.

Embryos for this study are collected from actively laying commercial turkeys. Embryo donor turkeys are inseminated weekly using pooled semen from male turkeys of the same breed to produce eggs for injection.

On the day of egg collection, fertile hens are euthanized 2h post oviposition by cervical dislocation. The recently ovulated and fertilized eggs are collected from the upper magnum region of the oviduct under sterile conditions and placed in a glass well and covered with Ringers' Medium and maintained at about 40°C until microinjection.

Cytoplasmic injection of artificial chromosomes containing the OMC24-IRES-LC is achieved using the microinjection apparatus disclosed in US Patent Application No. 09/919,143. Approximately 500 chromosomes are injected into the Stage I embryos at a single site.

Following microinjection, the embryos are transferred to the oviduct of recipient turkeys essentially as described in Olsen et al, B. J Exp Zool 109: 355-66, 1948. Typically, about one day after OT, the recipient turkeys lay a hard shell egg containing the manipulated ovum. Eggs are incubated in an incubator until hatching of the birds.

G2 transchromosomal turkeys are obtained which contain the artificial chromosome in their genome. The artificial chromosome containing the OMC24-IRES-HC is introduced into embryos obtained from the G2 turkeys in essentially the same manner as described for the OMC24-IRES-LC.

Eggs from G1 transchromosomal turkeys which contain both the OMC-IRES-LC and OMC24-IRES-HC containing chromosomes in their genome are tested for the presence of intact functional monoclonal antibody. A Costar flat 96-well plate is coated with 100 µl of C Goat-anti-Human kappa at a concentration of 5 µg/ml in PBS. The plate is incubated at 37 °C for two hours. 200 µl of 5% PBA is added to the wells followed by an incubation at 37 °C for about 60-90 minutes followed by a wash. 100 µl of egg white samples (diluted in 1% PBA:LBP) is added to each well and the plate is incubated at 37 °C for about 60-90 min followed by a wash. 100 µl of a 1:2000 dilution of F'2 Goat anti-Human IgG Fc-AP in 1% PBA is added to the wells and the plate is incubated at 37 °C for 60-90 min followed by a wash. The antibody is detected by placing 75 µl of 1mg/ml PNPP (p-nitrophenyl phosphate) in 5x developing buffer in each well and incubating for about 10-30 mins at room temperature. The detection reaction is stopped using 75ul of 1N NaOH. The egg white tests positive for significant levels of the antibody.

Example 25: Production of Interferon Using Avian Artificial Chromosomes in Quail

Artificial chromosomes comprising a chicken (Barred-Rock) chromosome centromere (CAC) are prepared essentially using methods described in US Patent No. 6,743,967, issued June 1, 2004, the disclosure of which is incorporated in its entirety herein by reference.

A coding sequence for interferon alpha 2b disclosed in US Patent Application No. 10/463,980, filed June 17, 2003, the disclosure of which is incorporated in its entirety herein by reference, is inserted in the 1OMC24-IRES1-Epo - ChromattB construct disclosed herein in Example 22 in place of the EPO
5 coding sequence to produce 1OMC24-IRES-INF - ChrommattB. The 1OMC24-IRES-INF - ChrommattB is cloned into the CACs essentially as described in Example 23. The recombinant CACs are prepared then isolated by a dual laser high-speed flow cytometer.

The flow-sorted chromosomes are pelleted by centrifugation and are diluted
10 to a concentration of about 10,000 chromosomes per μ l of injection buffer. Approximately 50 nanoliters (nl) of injection buffer is injected per quail embryo.

Embryos for this study are collected from actively laying quail. Embryo donor quail are inseminated weekly using pooled semen from male quail of the same breed to produce eggs for injection.

15 On the day of egg collection, fertile quail are euthanized 2h post oviposition by cervical dislocation. The recently ovulated and fertilized eggs are collected from the upper magnum region of the oviduct under sterile conditions and placed in a glass well and covered with Ringers' Medium and maintained at about 40°C until microinjection.

20 Cytoplasmic injection of artificial chromosomes is achieved using the microinjection apparatus disclosed in US Patent Application No. 09/919,143, filed July 31, 2001. Chromosomes are injected into the Stage I embryos at a single site in each embryo.

Following microinjection, the embryos are transferred to the oviduct of
25 recipient quail essentially as described in Olsen et al, B. J Exp Zool 109: 355-66, 1948. Typically, about one day after OT, the recipient quail lay a hard shell egg containing the manipulated ovum. Eggs are incubated in an incubator until hatching of the birds.

Eggs from G2 transchromosomal quail test positive for the presence of
30 intact functional interferon alpha 2b.

**Example 26: Production of Monoclonal Antibody Using Avian Artificial
Chromosomes in Chicken**

An attB site and a hygromycin^R coding sequence are inserted into the OMC24-IRES-LC and OMC24-IRES-HC vectors disclosed in US Patent
5 Application No. 10/856,218, filed July 31, 2001, which are then each cloned into CACs of Example 25 essentially as described in Examples 22 and 23. The CACs are isolated by a dual laser high-speed flow cytometer.

The flow-sorted chromosomes are pelleted by centrifugation and are diluted to a concentration of 7000-12,000 chromosomes per μ l of injection buffer.
10 Approximately 50 nanoliters (nl) of injection buffer is injected per chicken embryo.

Embryos for this study are collected from actively laying *G. gallus*. Embryo donor chickens are inseminated weekly using pooled semen from male chickens of the same breed to produce eggs for injection.

On the day of egg collection, fertile hens are euthanized 2h post oviposition
15 by cervical dislocation. The recently ovulated and fertilized eggs are collected from the upper magnum region of the oviduct under sterile conditions and placed in a glass well and covered with Ringers' Medium and maintained at about 41°C until microinjection.

Cytoplasmic injection of artificial chromosomes containing the OMC24-IRES-LC is achieved using the microinjection apparatus disclosed US Patent
20 Application No. 09/919,143. Approximately 500 chromosomes are injected into the Stage I embryos at a single site.

Following microinjection, the embryos are transferred to the oviduct of recipient chickens essentially as described in Olsen et al, B. J Exp Zool 109: 355-
25 66, 1948. Typically, about one day after OT, the recipient chickens lay a hard shell egg containing the manipulated ovum. Eggs are incubated in an incubator until hatching of the G0 birds.

G2 transchromosomal chickens are obtained which contain the artificial chromosome in their genome. The artificial chromosome containing the OMC24-IRES-HC is introduced into embryos obtained from the G2 chickens in essentially
30 the same manner as described for the OMC24-IRES-LC.

Eggs from G1 transchromosomal chickens which contain both the OMC-IRES-LC and OMC24-IRES-HC in their genome are tested for the presence of intact functional monoclonal antibody. A Costar flat 96-well plate is coated with 100 μ l of C Goat-anti-Human kappa at a concentration of 5 μ g/ml in PBS. The
5 plate is incubated at 37 °C for two hours. 200 μ l of 5% PBA is added to the wells followed by an incubation at 37 °C for about 60-90 minutes followed by a wash. 100 μ l of egg white samples (diluted in 1% PBA:LBP) is added to each well and the plate is incubated at 37 °C for about 60-90 min followed by a wash. 100 μ l of a 1:2000 dilution of F'2 Goat anti-Human IgG Fc-AP in 1% PBA is added to the
10 wells and the plate is incubated at 37 °C for 60-90 min followed by a wash. The antibody is detected by placing 75 μ l of 1mg/ml PNPP (p-nitrophenyl phosphate) in 5x developing buffer in each well and incubating for about 10-30 mins at room temperature. The detection reaction is stopped using 75 μ l of 1N NaOH. The egg white tests positive for significant levels of the antibody.

15

While this invention has been described with respect to various specific examples and embodiments, it is to be understood that the invention is not limited thereto and that it can be variously practiced with the scope of the following claims.

What is claimed is:

1. A transchromosomic avian.
- 5 2. The transchromosomic avian of claim 1 wherein the avian is a G1 transchromosomic avian.
3. The transchromosomic avian of claim 1 wherein the avian is a G2 transchromosomic avian.
- 10 4. The transchromosomic avian of claim 1 wherein the avian is a germline transchromosomic avian.
5. The transchromosomic avian of claim 1 wherein the avian is
15 selected from the group consisting of chicken, quail and turkey.
6. The transchromosomic avian of claim 1 wherein the artificial chromosome comprises a centromere.
- 20 7. The transchromosomic avian of claim 6 wherein the centromere is an insect centromere.
8. The transchromosomic avian of claim 6 wherein the centromere is a mammalian centromere.
- 25 9. The transchromosomic avian of claim 6 wherein the centromere is an avian centromere.
10. The transchromosomic avian of claim 1 wherein the artificial
30 chromosome comprises a heterologous nucleotide sequence.

11. The transchromosomic avian of claim 10 wherein the nucleotide sequence is heterologous to the avian.

12. The transchromosomic avian of claim 10 wherein the nucleotide
5 sequence is heterologous to the chromosome.

13. The transchromosomic avian of claim 10 wherein the heterologous nucleotide sequence includes a coding sequence for a therapeutic substance.

10 14. The transchromosomic avian of claim 10 wherein the heterologous nucleotide sequence includes a gene expression controlling region.

15 15. The transchromsomic avian of claim 14 wherein the gene expression controlling region is a promoter which is operable in a cell of an oviduct.

16. The transchromosomic avian of claim 14 wherein the gene expression controlling region includes a promoter selected from the group consisting of a lysozyme promoter, an ovomucin promoter, an ovomucoid
20 promoter and an ovalbumin promoter.

17. The transchromosomic avian of claim 1 wherein the avian lays an egg comprising a heterologous protein.

25 18. A method of producing a transchromosomic avian comprising:
inserting a heterologous nucleotide sequence in a chromosome;
substantially purifying the chromosome;
introducing the chromosome into an avian embryo;
maintaining the embryo under conditions suitable for the embryo to develop
30 and hatch as a chick;
thereby producing a transgenic avian.

19. The method of claim 18 wherein the avian can produce transchromosomic offspring.

20. The method of claim 18 wherein the chromosome is introduced by
5 microinjection.

21. The method of claim 18 wherein the chromosome is introduced into the embryo by delivering the chromosome to a germinal disc.

10 22. The method of claim 18 wherein the embryo is an early stage embryo.

23. The method of claim 18 wherein the embryo is a stage I embryo.

15 24. The method of claim 18 wherein between 1 and about 10,000 chromosomes are introduced into the embryo.

25. The method of claim 18 wherein between 1 and about 1,000 chromosomes are introduced into the embryo.

20

26. The method of claim 18 wherein the avian is selected from the group consisting of chicken, quail and turkey.

27. The method of claim 18 wherein the chromosome comprises a
25 heterologous nucleotide sequence.

28. The method of claim 27 wherein the heterologous nucleotide sequence includes a coding sequence for a therapeutic substance.

30 29. The method of claim 27 wherein the heterologous nucleotide sequence includes a gene expression controlling region.

30. The method of claim 29 wherein the gene expression controlling region includes a promoter selected from the group consisting of a lysozyme promoter, an ovomucin promoter, an ovomucoid promoter and an ovalbumin promoter.

5

31. An avian cell comprising an artificial chromosome wherein the artificial chromosome includes a nucleotide sequence encoding a therapeutic substance.

10

32. The avian cell of claim 31 wherein the avian is selected from the group consisting of chicken, quail and turkey.

33. The avian cell of claim 31 wherein the heterologous nucleotide sequence includes a gene expression controlling region.

15

34. A method of producing a transgenic vertebrate animal comprising:

introducing into an embryo of a vertebrate animal a recombination site such that the recombination site is present in ovum of a mature vertebrate animal developed from the embryo wherein the embryo does not normally comprise the recombination site,

20

introducing into the embryo of the vertebrate animal or fertilized ovum of the mature vertebrate animal a nucleotide sequence comprising a second recombination site and a coding sequence;

25

introducing into the embryo of the vertebrate animal or fertilized ovum of the mature vertebrate animal a substance which facilitates insertion of the nucleotide sequence comprising a second recombination site and a coding sequence proximal to the recombination site; and

30

exposing the embryo of the vertebrate animal or fertilized ovum of the mature vertebrate animal to conditions which lead to development of a viable transgenic vertebrate animal,

5 thereby producing a transgenic vertebrate animal.

35. The method of claim 34 wherein the vertebrate animal is an avian.

36. The method of claim 34 wherein the vertebrate animal is an avian
10 selected from the group consisting of chicken, turkey, duck, goose, quail, pheasants, parrots, finches, hawks, crows and ratites including ostrich, emu and cassowary.

37. The method of claim 34 wherein the vertebrate animal is a chicken.
15

38. The method of claim 34 wherein the embryo or fertilized ovum is selected from the group consisting of a stage I, stage II, stage III, stage IV, stage V, stage VI, stage VII, stage VIII, stage IX, stage X, stage XI and stage XII embryo.

20 39. The method of claim 34 wherein at least one of the recombination site or the nucleotide sequence comprising a second recombination site and a coding sequence is introduced into the embryo or fertilized ovum by transfection.

40. The method of claim 34 wherein at least one of the recombination
25 site or the nucleotide sequence comprising a second recombination site and a coding sequence is introduced into the embryo or fertilized ovum by microinjection.

41. The method of claim 34 wherein at least one of the recombination
30 site or the nucleotide sequence comprising a second recombination site and a coding sequence is introduced into the embryo or fertilized ovum by cell fusion.

42. The method of claim 34 wherein at least one of the recombination site or the nucleotide sequence comprising a second recombination site and a coding sequence is introduced into the embryo or fertilized ovum by lipofection.

5 43. The method of claim 34 wherein the recombination site or nucleotide sequence comprising a second recombination site and a coding sequence is introduced into the embryo or fertilized ovum in the presence of PEI.

44. The method of claim 34 wherein at least one of the recombination
10 site or the second recombination site is isolated from a bacteriophage.

45. The method of claim 34 wherein the recombination site is an attP site or an attB site.

15 46. The method of claim 34 wherein the second recombination site is an attP site or an attB site.

47. The method of claim 34 wherein the nucleotide sequence is stably
20 incorporated into the genome of the embryo or fertilized ovum.

48. The method of claim 34 wherein the substance is an enzyme.

49. The method of claim 48 wherein the enzyme is a site specific
25 recombinase.

50. The method of claim 48 wherein the enzyme is selected from the group consisting of a serine recombinase and a tyrosine recombinase.

51. The method of claim 48 wherein the enzyme is selected from the
30 group consisting of EcoYBCK, Φ C31, SCH10.38c, SCC88.14, SC8F4.15c, SCD12A.23, Bxb1, WwK, Sau CcrB, Bsu CisB, TP901-1, Φ 370.1, Φ 105, Φ FC1,

A118, Cac1956, Cac1951, Sau CcrA, Spn, TnpX, TndX, SPBc2, SC3C8.24, SC2E1.37, SCD78.04c, R4, Φ Rv1, Y4bA and Bja serine recombinases.

52. The method of claim 48 wherein the enzyme is selected from the
5 group consisting of Φ C31, TP901-1 and R4 serine recombinases.

53. The method of claim 48 wherein the enzyme is Φ C31 serine recombinase.

10 54. The method of claim 34 wherein the substance is nucleic acid.

55. The method of claim 54 wherein the nucleic acid is DNA or RNA.

56. The method of claim 55 wherein a vector comprises the DNA.
15

57. The method of claim 55 wherein the nucleic acid encodes an enzyme.

58. The method of claim 55 wherein the nucleic acid encodes a site
20 specific recombinase.

59. The method of claim 55 wherein the nucleic acid encodes an enzyme selected from the group consisting of a serine recombinase and a tyrosine recombinase.

25 60. The method of claim 55 wherein the nucleic acid encodes a serine recombinase selected from the group consisting of EcoYBCK, Φ C31, SCH10.38c, SCC88.14, SC8F4.15c, SCD12A.23, Bxb1, WwK, Sau CcrB, Bsu CisB, TP901-1, Φ 370.1, Φ 105, Φ FC1, A118, Cac1956, Cac1951, Sau CcrA, Spn, TnpX, TndX, SPBc2, SC3C8.24, SC2E1.37, SCD78.04c, R4, Φ Rv1, Y4bA and Bja serine recombinases.
30

61. The method of claim 55 wherein the nucleic acid encodes a serine recombinase selected from the group consisting of Φ C31, TP901-1 and R4 serine recombinases.

5 62. The method of claim 55 wherein the nucleic acid encodes a Φ C31 serine recombinase.

63. The method of claim 34 wherein the nucleotide sequence comprises an expression cassette.

10

64. The method of claim 34 wherein the coding sequence encodes a therapeutic composition.

65. The method of claim 64 wherein the therapeutic composition
15 comprises at least one of a light chain or a heavy chain of an antibody.

66. The method of claim 64 wherein the therapeutic composition is a cytokine.

20 67. The method of claim 64 wherein the therapeutic composition is selected from the group consisting of interferon, erythropoietin, and granulocyte-colony stimulating factor.

68. The method of claim 34 wherein the coding sequence encodes a
25 polypeptide present in an egg produced by the transgenic vertebrate animal wherein the animal is an avian.

69. An egg of claim 68.

30 70. The method of claim 34 wherein the recombination site is introduced into the embryo of a vertebrate animal or fertilized ovum before fertilization.

71. An ovum or sperm produced by a transgenic vertebrate animal of claim 34.
- 5 72. A transgenic vertebrate animal produced according to the method of claim 34.
73. A descendent of the transgenic vertebrate animal of claim 34.
74. A method of producing a transgenic vertebrate animal comprising:
- 10 introducing into an embryo of a vertebrate animal a recombination site such that the recombination site is present in sperm of a mature vertebrate animal developed from the embryo wherein the embryo does not normally comprise the recombination site,
- 15 fertilizing an ovum with sperm comprising the recombination site;
- introducing into the ovum a nucleotide sequence comprising a second recombination site and a coding sequence and a substance which facilitates
- 20 insertion of the nucleotide sequence comprising a second recombination site and coding sequence proximal to the recombination site; and
- exposing the fertilized ovum to conditions which lead to development of a viable transgenic vertebrate animal,
- 25 thereby producing a transgenic vertebrate animal.
75. The method of claim 74 wherein the vertebrate animal is an avian.
- 30 76. The method of claim 74 wherein the vertebrate animal is an avian selected from the group consisting of chicken, turkey, duck, goose, quail,

pheasants, parrots, finches, hawks, crows and ratites including ostrich, emu and cassowary.

5 77. The method of claim 74 wherein the vertebrate animal is a chicken.

 78. The method of claim 74 wherein the embryo is selected from the group consisting of a stage I, stage II, stage III, stage IV, stage V, stage VI, stage VII, stage VIII, stage IX, stage X, stage XI and stage XII embryo.

10 79. The method of claim 74 wherein at least one of the recombination site or the nucleotide sequence comprising a second recombination site and a coding sequence is introduced into the embryo or fertilized ovum by cell fusion.

 80. The method of claim 74 wherein at least one of the recombination
15 site or the nucleotide sequence comprising a second recombination site and a coding sequence is introduced into the embryo or fertilized ovum by lipofection.

 81. The method of claim 74 wherein at least one of the recombination
20 site or the nucleotide sequence comprising a second recombination site and a coding sequence is introduced into the embryo or fertilized ovum by transfection.

 82. The method of claim 74 wherein at least one of the recombination
25 site or the nucleotide sequence comprising a second recombination site and a coding sequence is introduced into the embryo or fertilized ovum by microinjection.

 83. The method of claim 74 wherein the recombination site or
30 nucleotide sequence comprising a second recombination site and a coding sequence is introduced in the presence of PEI.

 84. The method of claim 74 wherein at least one of the recombination site or the second recombination site is isolated from a bacteriophage.

85. The method of claim 74 wherein the recombination site is an attP site or an attB site.
- 5 86. The method of claim 74 wherein the second recombination site is an attP site or an attB site.
87. The method of claim 74 wherein the nucleotide sequence is stably incorporated into the genome of the embryo or fertilized ovum.
- 10 88. The method of claim 74 wherein the substance is an enzyme.
89. The method of claim 88 wherein the enzyme is a site specific recombinase.
- 15 90. The method of claim 88 wherein the enzyme is selected from the group consisting of a serine recombinase and a tyrosine recombinase.
91. The method of claim 88 wherein the enzyme is selected from the group consisting of EcoYBCK, Φ C31, SCH10.38c, SCC88.14, SC8F4.15c, SCD12A.23, Bxb1, WwK, Sau CcrB, Bsu CisB, TP901-1, Φ 370.1, Φ 105, Φ FC1, A118, Cac1956, Cac1951, Sau CcrA, Spn, TnpX, TndX, SPBc2, SC3C8.24, SC2E1.37, SCD78.04c, R4, Φ Rv1, Y4bA and Bja serine recombinases.
- 20 92. The method of claim 88 wherein the enzyme is selected from the group consisting of Φ C31, TP901-1 and R4 serine recombinases.
93. The method of claim 88 wherein the enzyme is Φ C31 serine recombinase.
- 25 94. The method of claim 74 wherein the substance is nucleic acid.
- 30

95. The method of claim 94 wherein the nucleic acid is DNA or RNA.

96. The method of claim 95 wherein a vector comprises the DNA.

5 97. The method of claim 94 wherein the nucleic acid encodes an enzyme.

98. The method of claim 94 wherein the nucleic acid encodes a site specific recombinase.
10

99. The method of claim 94 wherein the nucleic acid encodes an enzyme selected from the group consisting of a serine recombinase and a tyrosine recombinase.

15 100. The method of claim 94 wherein the nucleic acid encodes a serine recombinase selected from the group consisting of EcoYBCK, Φ C31, SCH10.38c, SCC88.14, SC8F4.15c, SCD12A.23, Bxb1, WwK, Sau CcrB, Bsu CisB, TP901-1, Φ 370.1, Φ 105, Φ FC1, A118, Cac1956, Cac1951, Sau CcrA, Spn, TnpX, TndX, SPBc2, SC3C8.24, SC2E1.37, SCD78.04c, R4, Φ Rv1, Y4bA and Bja serine
20 recombinases.

101. The method of claim 94 wherein the nucleic acid encodes a serine recombinase selected from the group consisting of Φ C31, TP901-1 and R4 serine recombinases.
25

102. The method of claim 94 wherein the nucleic acid encodes a Φ C31 serine recombinase.

103. The method of claim 74 wherein the nucleotide sequence comprises
30 an expression cassette.

104. The method of claim 74 wherein the coding sequence encodes a therapeutic composition.

105. The method of claim 104 wherein the therapeutic composition
5 comprises at least one of a light chain or a heavy chain of an antibody.

106. The method of claim 105 wherein the antibody is a human antibody.

107. The method of claim 104 wherein the therapeutic composition is a
10 cytokine.

108. The method of claim 104 wherein the therapeutic composition is
selected from the group consisting of interferon, erythropoietin, and granulocyte-
colony stimulating factor.

15

109. The method of claim 74 wherein the coding sequence encodes a
polypeptide present in an egg produced by the transgenic animal wherein the
animal is an avian.

20 110. An egg of claim 109.

111. The method of claim 74 wherein the recombination site is
introduced into the embryo of a vertebrate animal or fertilized ovum before
fertilization.

25

112. An ovum or sperm produced by a transgenic vertebrate animal of
claim 74.

113. A transgenic vertebrate animal produced according to the method of
30 claim 74.

114. A descendent of the transgenic vertebrate animal of claim 74.

115. A method of modifying a vertebrate cell genome comprising:

5 introducing into a vertebrate cell a recombination site such that the recombination site is inserted into the vertebrate cell genome wherein the genome does not normally comprise the recombination site;

10 introducing a nucleotide sequence comprising a second recombination site and a coding sequence into the vertebrate cell or progeny cell thereof; and

introducing into the vertebrate cell or progeny cell thereof a substance which facilitates insertion of the nucleotide sequence comprising a second recombination site and a coding sequence proximal to the recombination site,
15 thereby modifying a vertebrate cell genome.

116. The method of claim 115 wherein the vertebrate cell is a cell of an avian.

20 117. The method of claim 115 wherein the vertebrate cell is a chicken cell.

118. The method of claim 115 wherein the vertebrate cell is a germ line cell.

25 119. The method of claim 115 wherein the vertebrate cell is an ovum or an embryo.

30 120. The method of claim 115 wherein the vertebrate cell an embryo cell and the embryo is selected from the group consisting of a stage I, stage II, stage III, stage IV, stage V, stage VI, stage VII, stage VIII, stage IX, stage X, stage XI and stage XII embryo.

121. The method of claim 115 wherein at least one of the recombination site or the nucleotide sequence comprising a second recombination site and a coding sequence is introduced into the vertebrate cell genome by a method selected
5 from the group consisting of cell fusion, lipofection, transfection and microinjection.

122. The method of claim 115 wherein the recombination site or nucleotide sequence comprising a second recombination site and a coding sequence
10 is introduced into the vertebrate cell genome in the presence of PEI.

123. The method of claim 115 wherein at least one of the recombination site or the second recombination site is isolated from a bacteriophage.

15 124. The method of claim 115 wherein the recombination site is an attP site or an attB site.

125. The method of claim 115 wherein the second recombination site is an attP site or an attB site.
20

126. The method of claim 115 wherein the nucleotide sequence is stably incorporated into the genome of the vertebrate cell.

127. The method of claim 115 wherein the substance is an enzyme.
25

128. The method of claim 127 wherein the enzyme is a site specific recombinase.

129. The method of claim 127 wherein the enzyme is selected from the
30 group consisting of a serine recombinase and a tyrosine recombinase.

130. The method of claim 127 wherein the enzyme is selected from the group consisting of EcoYBCK, Φ C31, SCH10.38c, SCC88.14, SC8F4.15c, SCD12A.23, Bxb1, WwK, Sau CcrB, Bsu CisB, TP901-1, Φ 370.1, Φ 105, Φ FC1, A118, Cac1956, Cac1951, Sau CcrA, Spn, TnpX, TndX, SPBc2, SC3C8.24,
5 SC2E1.37, SCD78.04c, R4, Φ Rv1, Y4bA and Bja serine recombinases.

131. The method of claim 127 wherein the enzyme is Φ C31 serine recombinase.

10 132. The method of claim 115 wherein the substance is nucleic acid.

133. The method of claim 132 wherein the nucleic acid encodes an enzyme.

15 134. The method of claim 132 wherein the nucleic acid encodes a site specific recombinase.

135. The method of claim 132 wherein the nucleic acid encodes an enzyme selected from the group consisting of a serine recombinase and a tyrosine
20 recombinase.

136. The method of claim 127 wherein the nucleic acid encodes a serine recombinase selected from the group consisting of EcoYBCK, Φ C31, SCH10.38c, SCC88.14, SC8F4.15c, SCD12A.23, Bxb1, WwK, Sau CcrB, Bsu CisB, TP901-1,
25 Φ 370.1, Φ 105, Φ FC1, A118, Cac1956, Cac1951, Sau CcrA, Spn, TnpX, TndX, SPBc2, SC3C8.24, SC2E1.37, SCD78.04c, R4, Φ Rv1, Y4bA and Bja serine recombinases.

137. The method of claim 127 wherein the nucleic acid encodes a Φ C31
30 serine recombinase.

138. The method of claim 115 wherein the coding sequence encodes a therapeutic composition.

139. The method of claim 138 wherein the therapeutic composition
5 comprises at least one of a light chain or a heavy chain of an antibody.

140. The method of claim 138 wherein the therapeutic composition is a cytokine.

10 141. The method of claim 138 wherein the therapeutic composition is selected from the group consisting of interferon, erythropoietin, and granulocyte-colony stimulating factor.

142. A cell produced according to the method of claim 115.
15

20

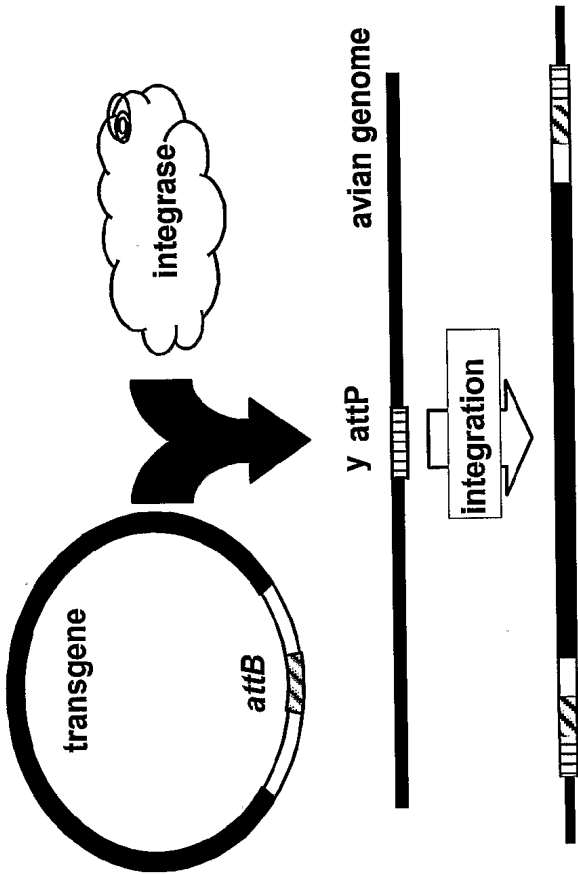


Fig. 1

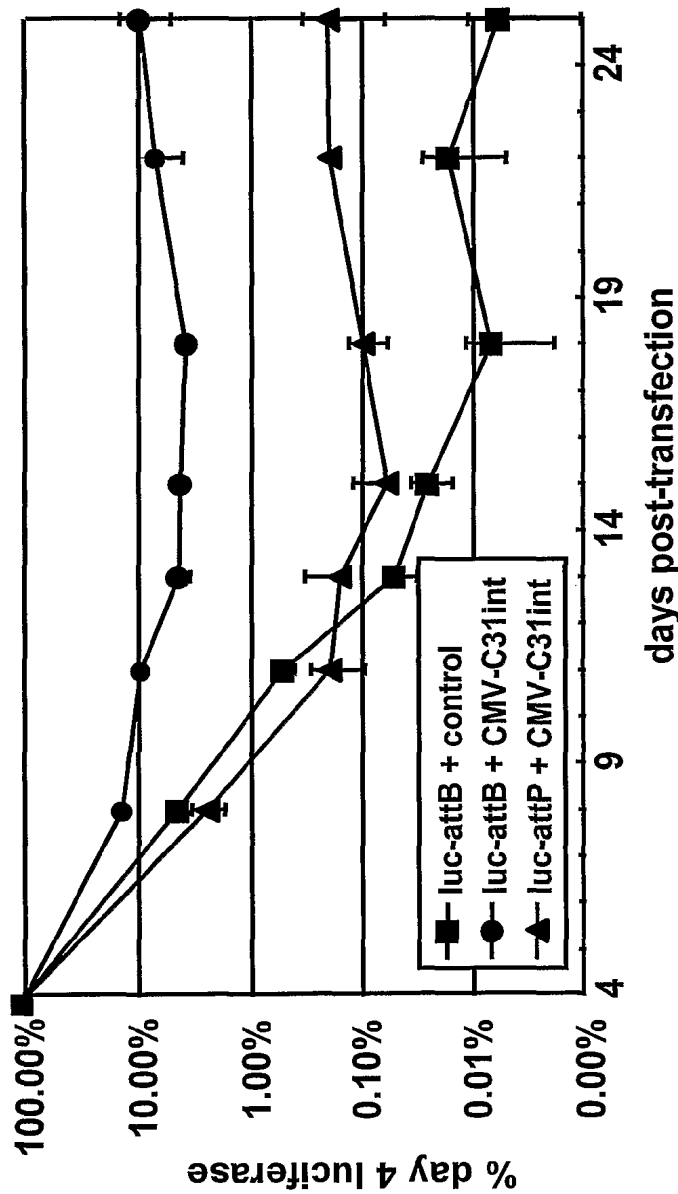


Fig. 2

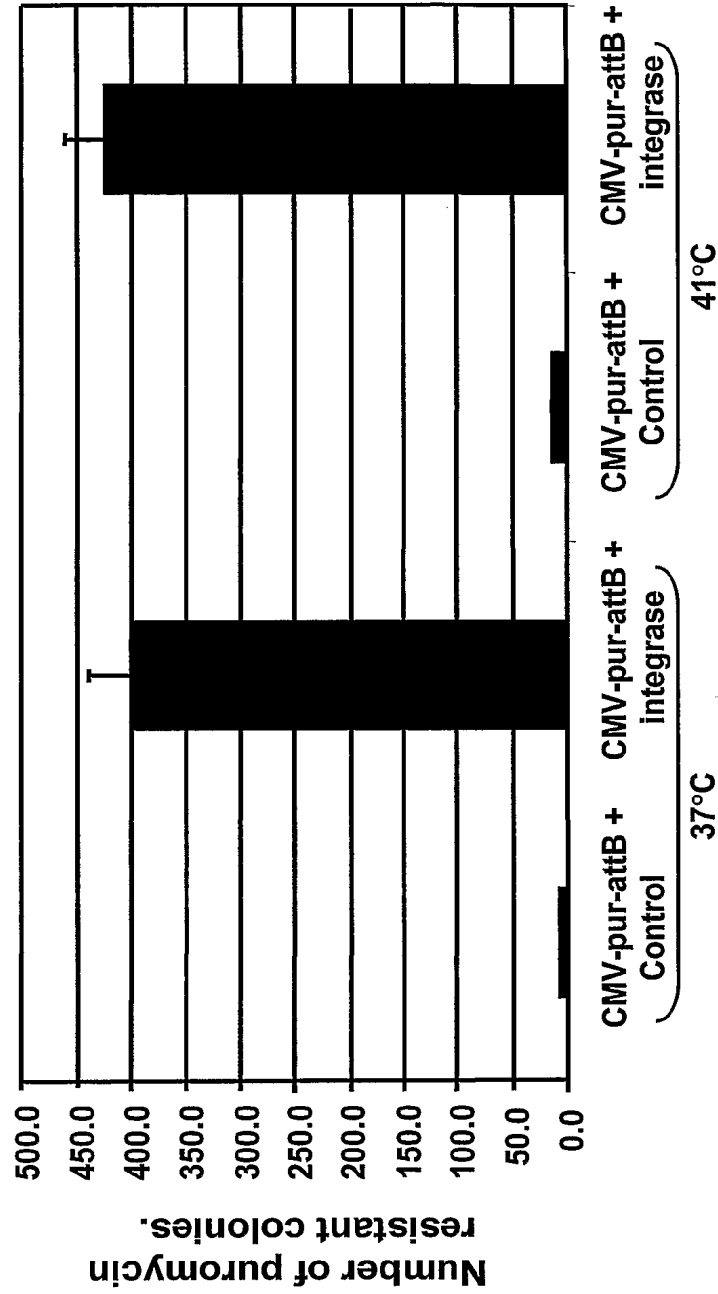


Fig. 3

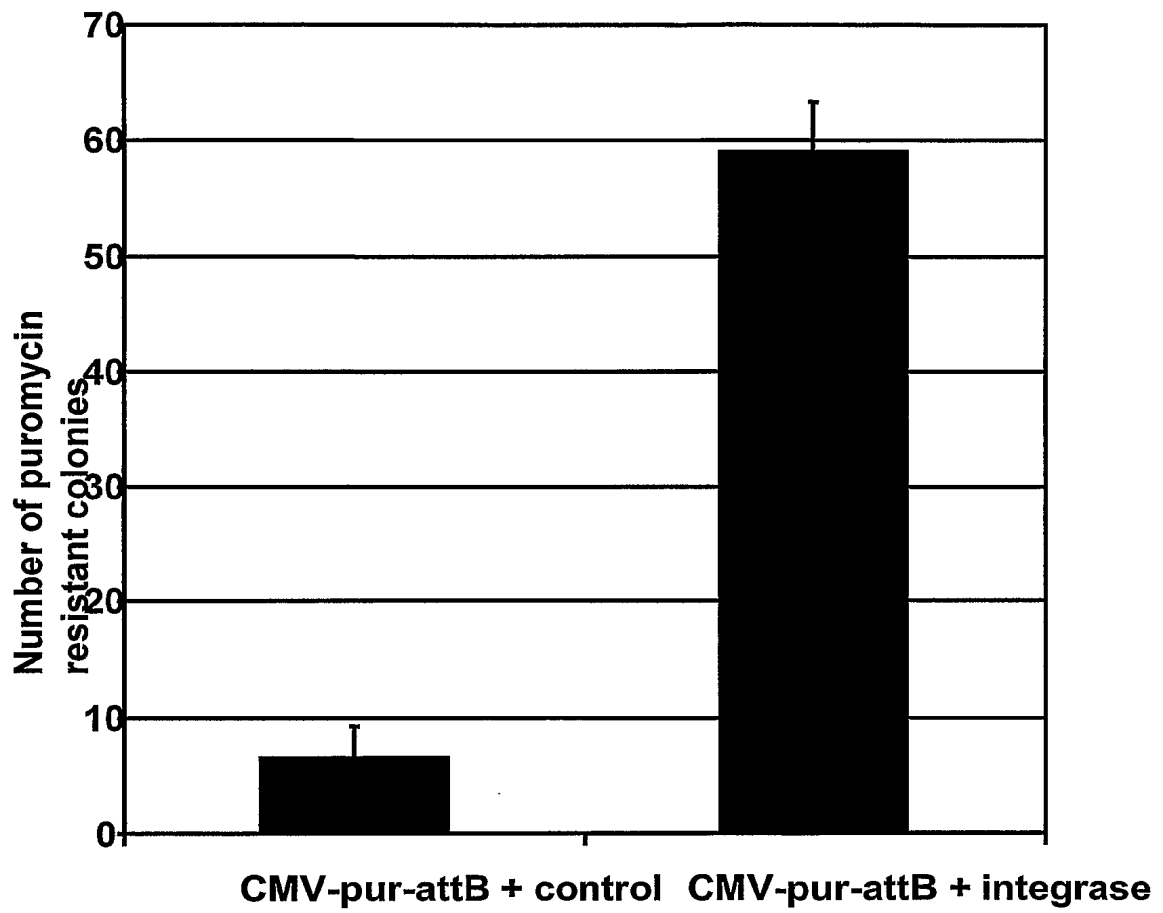


Fig. 4

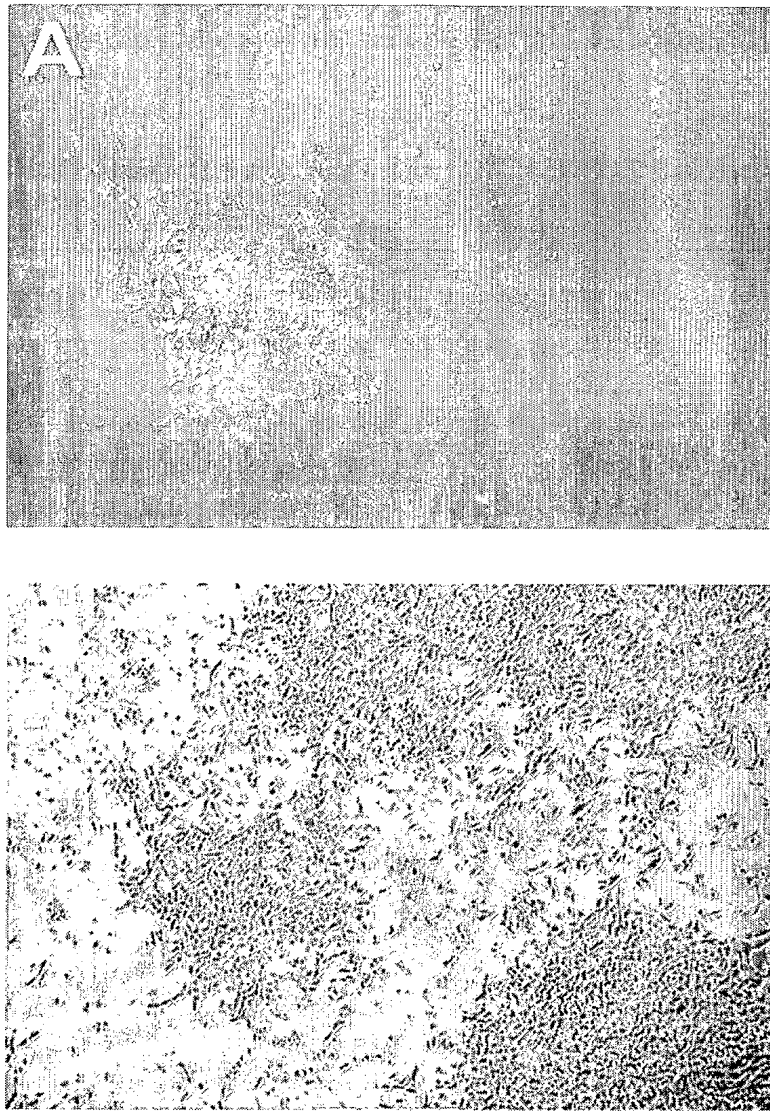
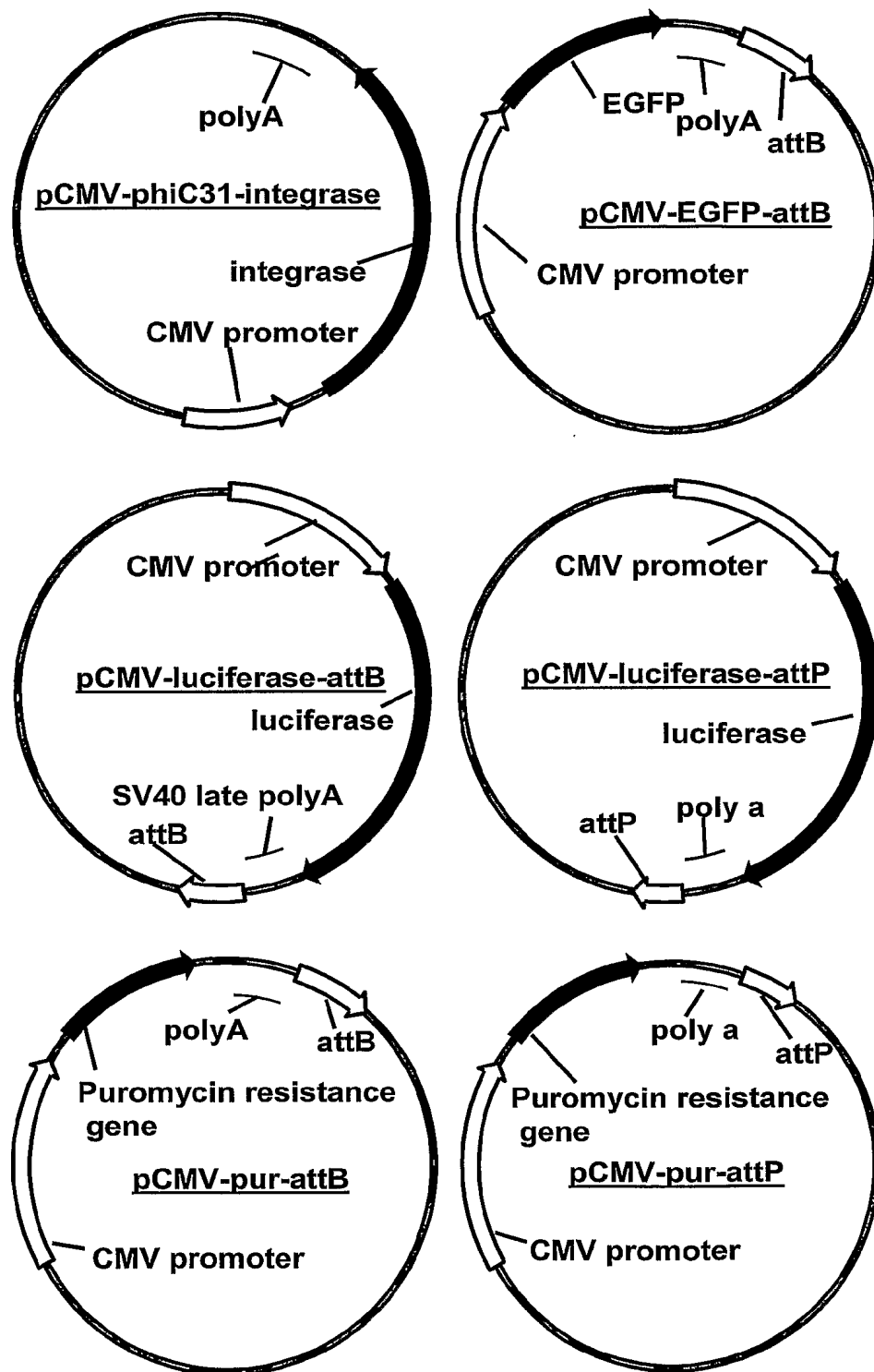
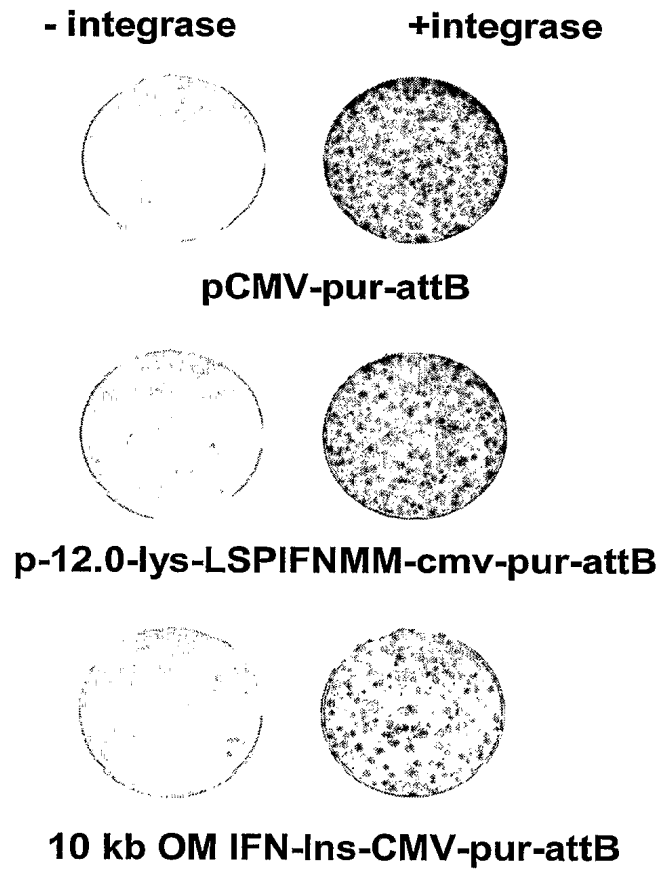
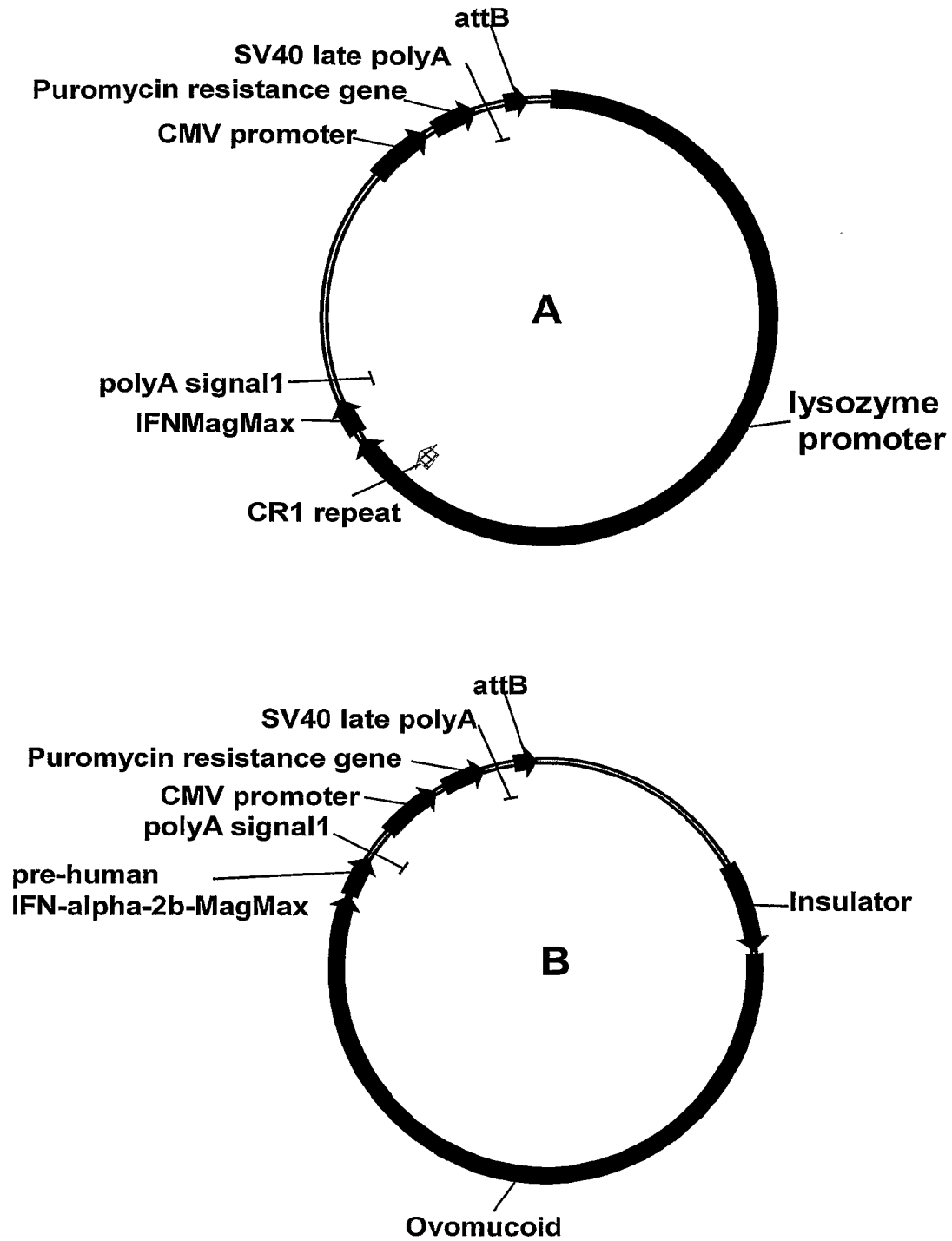


Fig. 5

**Fig. 6**

***Fig. 7***

**Fig. 8**

pCMV-C31int (SEQ ID NO: 1)

CATTTCGCCATTTCAGGCTGCGCAACTGTTGGGAAGGGCGATCGGTGCGGGCCTCTTCGCTATT
ACGCCAGCCAATACGCAAACCGCCTCTCCCCGCGCGTTGGCCGATTCAATTAATGCAGGATCG
ATCCAGACATGATAAGATACATTGATGAGTTTGGACAAACCACAACCTAGAATGCAGTGAAAA
AAATGCTTTTATTTGTGAAATTTGTGATGCTATTGCTTTATTTGTAAACATTATAAGCTGCAA
TAAACAAGTTAACAACAACAATTGCATTCATTTTATGTTTCAGGTTTCAGGGGGAGGTGTGGG
AGGTTTTTTTAAAGCAAGTAAAACCTCTACAAATGTGGTATGGCTGATTATGATCATGAACAG
ACTGTGAGGACTGAGGGGCCCTGAAATGAGCCTTGGGACTGTGAATCTAAAATACACAAACAA
TTAGAATCACTAGCTCCTGTGTATAATATTTTCATAAATCATACTCAGTAAGCAAAACTCTC
AAGCAGCAAGCATATGCAGCTAGTTTAAACACATTATACACTTAAAAATTTTATATTTACCTT
AGAGCTTTAAATCTCTGTAGGTAGTTTGTCCAATTATGTCACACCACAGAAGTAAGGTTCCCT
TCACAAAGATCCCAAGCTAGCTTATAATACGACTCACTATAGGGAGAGAGCTATGACGTCGC
ATGCACGCGTAAGCTTGGGCCCTCGAGGGATCCGGGTGTCTCGCTACGCCGCTACGCTCTC
CGTGCCGTCCTGGGCGTCGTCTTCGTGCTCGTCGGTCGGCGGCTTCGCCCACGTGATCGAAG
CGCGCTTCTCGATGGGCGTTCCTTGCCCCCTGCCGCTAGTCGACTTCGTGACAACGATCTTG
TCTACGAAGAGCCGACGAACACGCGCTTGTCTACTGACGCGCGCCCCACACGACTT
AGGCCGCGTCGGGTACGCGTCGGCGTCTTCGGGGAACCATTTGGTCAAGGGGAAGCTTCGGGG
CTTCGGCGGCTTCAAGTTCGGCAAGCCGCTCTTCGCCCTTGCTGCCGGAGCGTCAGCGCT
GCCTGTTGCTTCCGGAAGTGCTTCCTGCCAACGGGTCCGTGTCACGCGCTGCCGCGCGGTC
TTCGTACAGCTCTTCAAGGGCGTTCAGGGCGTCGGCGCGCTCCGCAACAAGGTTCCGCCGTT
CGCCGCTCTTCTCAGGCGCTCAGTGAGCTTGCCGAAGCGTCGGGCGGCTTCCCACAGAAGC
GCCAACGCTCTCTTCGTGCGCTTCGGCGTGCTGATCTTGTGTAAGATGCGTTCCGCAACGAA
CTTGTGAGTGC CGCCATGCTGACGTTGACGTGCTTTCGTGCTGCCAGGTGCCGAGCGGGT
CGACCACCTTCCGGCGACGGCAGCGGTAAGAGTCTTGATCGATTCTTCCCCGCGCTTCGAA
GTCATGACGGCGCCACACTCGCAGTACAGCTTGTCCATGGCGGACAGAATGGCTTGCCCCCG
GGAAAGCCCCCTTGCCGCGCCCCCTGCCGTCCAACCACGCTGAAGCTCATACCACTCAGCGG
GCTCGATGATCGGTCCGCAATCAAGCTCGACCGGCCGGAGCGTGATCGGGTCGCGCTGAATG
CGGTAACCCCTCAATCTTCGTGGTCGGCGTGCCGTCCGGCTTCTTCTTGATGATCACCTCAGC
GGCGAAGCCCGCAATACGCGGGTCCCGAAGGATTGCGATAACGGTTGCCGGGTCCCAGGCGC
TTGAAGCGGTCTTCTTCCCAATCGTCTCGCCCCGGTTCGGCACGGCGTCAGCGTCCATGCGC
TTACAAAGCCCCGTGATGCTGCCCGGGTGAATGGCGGCTTGACTGCCCGGCTTGAAGGGAAG
GTGTTTGTGCGTCTTGATCTCACGCCACCACCACGGATTACGTGGGCTCGAACTCGAAGG
GTCCGGTAAGGGGAGTGGTCGAGTGC GCAAGCTTGTGATGACGACATTGACCATTCGGCCG
TTGCGCGTGATCTCCTTCGTCTCCGAAACAAGCTCGAAGCCGTAAGGCGCCTTCCCCGCCGAC
GTACCCGCCCAATTTCGCGCTGAAGGTTCTTCGTGTCGAGAATCTTCGCCGACTTCAGCGAAG
ATTCCTTTGTGCGACGCGTCGAGCCGCATAATCAGGTGAATCAGGTCCATGACGTTTCCCTGC
CGGAAGACGCCTTCTGAGTGGAACAATCGTCACGCCAGGGCGAGCAATTCGAGACAAT
CGGAATCGCGTCCATGACCTTCAGGCGCGAGAAGCGCGACACGTCATAGACAATGATCATGT
TGAGCCGCCCGCGCGGCATTTCGTTCAAGATGCGTTCGAACTCCGGGCGCTCCGCCGTCCCG
AACGCCGACGTGCCCGGCGCTTCGTGAAATGCCCGACGAACCTGAACCGGCCCGCTCGCG
CTCGACTTCGCGCTGAAGGTGGGCCCTTGTCTTCGTTGGCGCTACGCTGTGTGCTGGGC
TTGCTGCGCTCGAATTCTCGCGCTCGCGCGACTGACGGTCGTAAGCACCCGCGTACGTGTCC
ACCCCGGTCAACAACCCCTTGTGTGTCATGTCGGCGACCCTACGACTAGTGAGCTCGTCGACCCG
GGAATTCGGGACCGGTACCTGCAGGCGTACCTTCTATAGTGTACCTAAATAGCTTTTTTGCA
AAAGCCTAGGCTAGAGTCCGGAGGCTGGATCGGTCCCGGTGTCTTCTATGGAGGTCAAAACA
GCGTGGATGGCGTCTCCAGGCGATCTGACGGTTCATAAACGAGCTCTGCTTATATAGACCT
CCCACCGTACACGCTACCGCCCATTTGCGTCAATGGGGCGGAGTTGTTACGACATTTTGGGA
AAGTCCCGTTGATTTTGGTGCCAAAACAACTCCATTGACGTCAATGGGGTGGAGACTTGG
AAATCCCCGTGAGTCAAACCGCTATCCACGCCCATTGATGTACTGCCAAAACCGCATCACCA
TGGTAATAGCGATGACTAATACGTAGATGTAATGCCAAGTAGGAAAGTCCATAAGGTTCATG
TACTGGGCATAATGCCAGGCGGGCCATTTACGTCATTGACGTCAATAGGGGCGTACTTGG
CATATGATACACTTGATGTACTGCCAAGTGGGCAGTTTACCGTAAATACTCCACCCATTGAC
GTCAATGGAAAGTCCCTATTGGCGTTACTATGGGAACATACGTCAATTATTGACGTCAATGGG
CGGGGGTTCGTTGGGCGGTACGCCAGGCGGGCCATTTACCGTAAGTTATGTAACGACCTGCAC

GATGCTGTTTCCTGTGTGAAATTGTTATCCGCTCACAATTCCACACATTATACGAGCCGGAA
GCTATAAAGTGTAAGCCTGGGGTGCCTAATGAGTGAAAGGGCCTCGTATACGCCTATTTTT
ATAGGTTAATGT CATGATAATAATGGTTTCTTAGACGTCAGGTGGCACTTTTTCGGGGAAATG
TGCGCGGAACCCCTATTTGTTTATTTTCTAAATACATTCAAATATGTATCCGCTCATGAGA
CAATAACCCCTGATAAATGCTTCAATAATATTGAAAAACGCGCGAATTGCAAGCTCTGCATTA
ATGAATCGGCCAACGCGCGGGGAGAGGCGGTTTTCGTATTGGGCGCTCTTCCGCTTCCTCGC
TCACTGACTCGCTGCGCTCGGTCTGCTCGGCTGCGGCGAGCGGTATCAGCTCACTCAAAGGCG
GTAATACGGTTATCCACAGAATCAGGGGATAACGCAGGAAAGAACATGTGAGCAAAAGGCCA
GCAAAAGGCCAGGAACCGTAAAAAGGCCGCGTTGCTGGCGTTTTTCCATAGGCTCCGCCCCC
CTGACGAGCATCACAAAAATCGACGCTCAAGTCAGAGGTGGCGAAACCCGACAGGACTATAA
AGATACCAGGCGTTTCCCCCTGGAAGCTCCCTCGTGCGCTCTCCTGTTCCGACCTGCCGCT
TACCGGATACCTGTCCGCTTTCTCCCTTCGGGAAGCGTGGCGCTTTCTCAATGCTCAGCGT
GTAGGTATCTCAGTTCGGTGTAGGTCTGCTCCAAGCTGGGCTGTGTGCACGAACCCCCC
GTTTCAGCCCCGACCGCTGCGCCTTATCCGGTAACCTATCGTCTTGAGTCCAACCCGGTAAGACA
CGACTTATCGCCACTGGCAGCAGCCACTGGTAACAGGATTAGCAGAGCGAGGTATGTAGGCG
GTGCTACAGAGTTCTTGAAGTGGTGGCCTAACTACGGCTACACTAGAAGGACAGTATTTGGT
ATCTGCGCTCTGCTGAAGCCAGTTACCTTCGGGAAAAGAGTTGGTAGCTCTTGATCCGGCAA
ACAAACCACCGCTGGTAGCGGTGGTTTTTTTTGTTTGCAAGCAGCAGATTACGCGCAGAAAAA
AAGGATCTCAAGAAGATCCTTTGATCTTTTTCTACGGGTCTGACGCTCAGTGGAAACGAAAA
TCACGTTAAGGGATTTTGGTCATGCCATAACTTCGTATAGCATACTTATACGAAGTTATGG
CATGAGATTATCAAAAAGGATCTTCACCTAGATCCTTTTAAATTAAAAATGAAGTTTTAAAT
CAATCTAAAGTATATATGAGTAAACTTGGTCTGACAGTTACCAATGCTTAATCAGTGAGGCA
CCTATCTCAGCGATCTGTCTATTTCTGTTTCATCCATAGTTGCCTGACTCCCCGTCGTGTAGAT
AACTACGATACGGGAGGGCTTACCATCTGGCCCCAGTGCTGCAATGATACCGCGAGACCCAC
GCTCACCGGCTCCAGATTTATCAGCAATAAACCCAGCCAGCCGGAAGGGCCGAGCGCAGAAGT
GGTCTGCAACTTTATCCGCTCCATCCAGTCTATTAATTGTTGCCGGGAAGCTAGAGTAAG
TAGTTCGCCAGTTAATAGTTTTCGCAACGTTGTTGCCATTGCTACAGGCATCGTGGTGTAC
GCTCGTCTGTTTGGTATGGCTTCATTTCAGCTCCGGTTCCCAACGATCAAGGCGAGTTACATGA
TCCCCCATGTTGTGCAAAAAAGCGGTTAGCTCCTTCGGTCTCCTCGATCGTTGTGCAAGTAA
GTTGGCCGAGTGTTATCACTCATGGTTATGGCAGCACTGCATAATTCTCTTACTGTCTATGC
CATCCGTAAGATGCTTTTCTGTGACTGTGTGAGTACTCAACCAAGTCATTCTGAGAATAGTGT
ATGCGGCGACCGAGTTGCTCTTGCCCGCGCTCAATACGGGATAATACCGCGCCACATAGCAG
AACTTTAAAAGTGCTCATCATTGGAACCGTTCTTCGGGGCGAAAACTCTCAAGGATCTTAC
CGCTGTTGAGATCCAGTTTCGATGTAACCCACTCGTGACCCCACTGATCTTCAGCATCTTTT
ACTTTCACCAGCGTTTCTGGGTGAGCAAAAACAGGAAGGCAAAATGCCGCAAAAAAGGGAAT
AAGGGCGACACGGAAATGTTGAATACTCATACTCTTCCTTTTTCAATATTATTGAAGCATTT
ATCAGGGTTATTGTCTCATGCCAGGGGTGGGCACACATATTTGATACCAGCGATCCCTACAC
AGCACATAATTCAATGCGACTTCCCTCTATCGCACATCTTAGACCTTTATTCTCCCTCAGC
ACACATCGAAGCTGCCGAGCAAGCCGTTCTCACCAGTCCAAGACCTGGCATGAGCGGATACA
TATTTGAATGTATTTAGAAAAATAAACAAATAGGGGTTCCGCGCACATTTCCCCGAAAAGTG
CCACCTGAAATTGTAAACGTTAATATTTTGTAAATTCGCGTTAAATTTTTGTAAATCAG
CTCATTTTTTAAACCAATAGGCCGAAATCGGCAAAATCCCTTATAAATCAAAAGAATAGACCG
AGATAGGGTTGAGTGTTGTTCCAGTTTGGAAACAAGAGTCCACTATTAAAGAACGTGGACTCC
AACGTCAAAGGGCGAAAAACCGTCTATCAGGGCGATGGCCCACTACGTGAACCATCACCCCTA
ATCAAGTTTTTTGGGGTCGAGGTGCCGTAAAGCACTAAATCGGAACCCCTAAAGGGAGCCCC
GATTTAGAGCTTGACGGGGAAAGCCGGCGAACGTGGCGAGAAAGGAAGGGAAGAAAGCGAAA
GGAGCGGGCGCTAGGGCGCTGGCAAGTGTAGCGGTACGCTGCGCGTAACCACACACCCGC
CGCGCTTAATGCGCCGCTACAGGGCGCGTC

Fig. 9

pCMV-luc-attB (SEQ ID NO: 2)

CTCTATCGATAGGTACCGAGCTCTTACGCGTGCTAGCCCTCGAGCAGGATCTATACATTGAA
TCAATATTGGCAATTAGCCATATTAGTCATTGGTTATATAGCATAAATCAATATTGGCTATT
GGCCATTGCATACGTTGTATCTATATCATAATATGTACATTTATATTGGCTCATGTCCAATA
TGACCGCCATGTTGACATTGATTATTGACTAGTTATTAATAGTAATCAATTACGGGGTCATT
AGTTCATAGCCCATATATGGAGTTCGCGGTTACATAAATTACGGTAAATGGCCCGCCTGGCT
GACCGCCCAACGACCCCCGCCCATTGACGTCAATAATGACGTATGTTCCCATAGTAACGCCA
ATAGGGACTTTCCATTGACGTCAATGGGTGGAGTATTTACGGTAAACTGCCCACTTGGCAGT
ACATCAAGTGTATCATATGCCAAGTCCGCCCCCTATTGACGTCAATGACGGTAAATGGCCCG
CCTGGCATTATGCCCAGTACATGACCTTACGGGACTTTCTACTTGGCAGTACATCTACGTA
TTAGTCATCGCTATTACCATGGTGATGCGGTTTTTGGCAGTACATCAATGGGCGTGGATAGCG
GTTTGACTCACGGGGATTTCGAAGTCTCCACCCCATTGACGTCAATGGGAGTTTGTTTTGGC
ACCAAATCAACGGGACTTTCCAAAATGTCTGTAACAACTCCGCCCCATTGACGCAATGGGC
GGTAGGCGTGTACGTTGGGAGGTCTATATAAGCAGAGCTCGTTTAGTGAACCGTCAGATCGC
CTGGAGACGCCATTCACGCTGTTTTGACCTCCATAGAGAAGACACCGGGACCGTCCAGCTCC
CCTCGAAGCTCGACTCTAGGGGCTCGAGATCTGCGATCTAAGTAAGCTTGGCATTCCGGTAC
TGTTGGTAAAGCCACCATGGAAGACGCCAAAAACATAAAGAAAGGCCCGCGCCATTCTATC
CGCTGGAAGATGGAACCGCTGGAGAGCAACTGCATAAGGCTATGAAGAGATACGCCCTGGTT
CCTGGAACAATTGCTTTTACAGATGCACATATCGAGGTGGACATCACTTACGCTGAGTACTT
CGAAATGTCCGTTTCGTTTGGCAGAAGCTATGAAACGATATGGGCTGAATACAAATCACAGAA
TCGTCGTATGCAGTGAAACTCTCTCAATTCTTTATGCCGGTGTGGGCGCGTTATTTATC
GGAGTTGCAGTTTGGCCCCGCGAACGACATTTATAATGAACGTGAATTGCTCAACAGTATGGG
CATTTTCGCAGCCTACCGTGGTGTTCGTTTCCAAAAAGGGGTTGCAAAAAATTTTGAACGTGC
AAAAAAGCTCCCAATCATCAAAAAATTATTATCATGGATTCTAAAACGGATTACCAGGGA
TTTCAGTCGATGTACACGTTTCGTACATCTCATCTACCTCCCGGTTTTTAATGAATACGATTT
TGTGCCAGAGTCCTTCGATAGGGACAAGACAATTGCACTGATCATGAACCTCTCTGGATCTA
CTGGTCTGCCTAAAGGTGTCTGCTCTGCCTCATAGAAGTGCCTGCGTGAGATTCTCGCATGCC
AGAGATCCTATTTTTTGGCAATCAAATCATTCCGGATACTGCGATTTTAAAGTGTGTTCATT
CCATCACGGTTTTTGGAAATGTTTACTACACTCGGATATTTGATATGTGGATTTTCGAGTCGTCT
TAATGTATAGATTTGAAGAAGAGCTGTTTCTGAGGAGCCTTCAGGATTACAAGATTCAAAGT
GCGCTGCTGGTGCCAAACCCTATTCTCCTTCTTCGCCAAAAGCACTCTGATTGACAAATACGA
TTTATCTAATTTACACGAAATTGCTTCTGGTGGCGCTCCCCTCTCTAAGGAAGTCCGGGAAG
CGGTTGCCAAGAGGTTCCATCTGCCAGGTATCAGGCAAGGATATGGGCTCACTGAGACTACA
TCAGCTATTCTGATTACACCCGAGGGGGATGATAAACCGGGCGCGGTCCGTAAAGTTGTTCC
ATTTTTTGAAGCGAAGGTTGTGGATCTGGATACCGGGAAAACGCTGGGCGTTAATCAAAGAG
GCGAACTGTGTGTGAGAGGTCCTATGATTTATGTCGGTTATGTAAACAATCCGGAAGCGACC
AACGCCTTGATTGACAAGGATGGATGGCTACATTCTGGAGACATAGCTTACTGGGACGAAGA
CGAACACTTCTTCATCGTTGACCGCCTGAAGTCTCTGATTAAGTACAAAGGCTATCAGGTGG
CTCCCGCTGAATTGGAATCCATCTTGCTCCAACACCCCAACATCTTCGACGCAGGTGTGCGA
GGTCTTCCCGACGATGACGCCGCTGAACCTCCCGCCGCGTTGTTGTTTTTGGAGCACGGA
GACGATGACGGAAGAGATCGTGGATTACGTGCGCAGTCAAGTAACAACCGCGAAGAAAGT
TGCGCGGAGGAGTTGTGTTTTGTGGACGAAGTACCGAAAGGCTTACCAGGAAAACCTCGACGCA
AGAAAAATCAGAGAGATCCTCATAAAGGCCAAGAAGGGCGGAAAGATCGCCGTGTAAATCTA
GAGTCGGGGCGGCCCGGCTTCGAGCAGACATGATAAGATACATTGATGAGTTTGGACAAA
CCACAACCTAGAATGCAGTGAAAAAATGCTTTATTTGTGAAATTTGTGATGCTATTGCTTTA
TTTGTAACCATTATAAGCTGCAATAAACAAGTTAACAACAACAATTGCATTCATTTTATGTT
TCAGGTTTCAGGGGGAGGTGTGGGAGGTTTTTTAAAGCAAGTAAACCTCTACAAATGTGGTA
AAATCGATAAGGATCAATTCCGGCTTCAGGTACCGTTCGACGATGTAGGTACAGGTCTCGAAGC
CCGGTTCGGGTGCCAGGGCGTGCCCTTGGGCTCCCGGGCGCGTACTCCACCTCACCCTATC
TGCTCCATGATGATGAACGGGTTCGAGGTGGCGGTAGTTGATCCCGGCGAACGCGCGCGCAC
CGGGAAGCCCTCGCCCTCGAAACCGCTGGGCGCGGTGGTCACGGTGAGCACGGGACGTGCGA
CGGCGTTCGGCGGTGCGGATACGCGGGGCAGCGTCAGCGGGTTCTCGACGGTCACGGCGGGC
ATGTCGACAGCCGAATTGATCCGTCGACCGATGCCCTTGAGAGCCTTCAACCCAGTCAGCTC
CTTCCGGTGGGCGCGGGGCATGACTATCGTCGCCGCACTTATGACTGTCTTCTTTATCATGC

AACTCGTAGGACAGGTGCCGGCAGCGCTCTTCCGCTTCCTCGCTCACTGACTCGCTGCGCTC
GGTCGTTTCGGCTGCGGCGAGCGGTATCAGCTCACTCAAAGGCGGTAATACGGTTATCCACAG
AATCAGGGGATAACGCAGGAAAGAACATGTGAGCAAAAGGCCAGCAAAGGCCAGGAACCGT
AAAAAGGCCGCGTTGCTGGCGTTTTTCCATAGGCTCCGCCCCCTGACGAGCATCACAAAAA
TCGACGCTCAAGTCAGAGGTGGCGAAACCCGACAGGACTATAAAGATACCAGGCGTTTTCCCC
CTGGAAGCTCCCTCGTGCGCTCTCCTGTTCCGACCCTGCCGCTTACCGGATACCTGTCCGCC
TTTCTCCCTTCGGGAAGCGTGGCGCTTTCTCAATGCTCACGCTGTAGGTATCTCAGTTCCGT
GTAGGTGCTTCGCTCCAAGCTGGGCTGTGTGCACGAACCCCCCGTTACGCCCCGACCGCTGCG
CCTTATCCGGTAACTATCGTCTTGAGTCCAACCCGGTAAGACACGACTTATCGCCACTGGCA
GCAGCCACTGGTAACAGGATTAGCAGAGCGAGGTATGTAGGCGGTGCTACAGAGTTCTTGAA
GTGGTGGCCTAACTACGGCTACACTAGAAAGGACAGTATTTGGTATCTGCGCTCTGCTGAAGC
CAGTTACCTTCGGAAAAAGAGTTGGTAGCTCTTGATCCGGCAAACAAACCACCGCTGGTAGC
GGTGGTTTTTTTTGTTTGCAAGCAGCAGATTACGCGCAGAAAAAAGGATCTCAAGAAGATCC
TTTGATCTTTTTCTACGGGGTCTGACGCTCAGTGGAAACGAAACTCACGTTAAGGGATTTTTGG
TCATGAGATTATCAAAAAGGATCTTCACCTAGATCCTTTTAAATTAATAAATGAAGTTTTAAA
TCAATCTAAAGTATATATGAGTAACTTGGTCTGACAGTTACCAATGCTTAATCAGTGAGGC
ACCTATCTCAGCGATCTGTCTATTTTCGTTCACTCCATAGTTGCCTGACTCCCCGTGCTGTAGA
TAACTACGATACGGGAGGGCTTACCATCTGGCCCCAGTGCTGCAATGATACCCGAGAGACCA
CGCTCACCGGCTCCAGATTTATCAGCAATAAACAGCCAGCCGGAAGGGCCGAGCGCAGAAAG
TGGTCCTGCAACTTTATCCGCCTCCATCCAGTCTATTAATTGTTGCCGGGAAGCTAGAGTAA
GTAGTTCCGCCAGTTAATAGTTTGCGCAACGTTGTTGCCATTGCTACAGGCATCGTGGTGTCA
CGCTCGTCGTTTTGGTATGGCTTCATTCAGCTCCGGTTCCTAACGATCAAGGCGAGTTACATG
ATCCCCCATGTTGTGCAAAAAAGCGGTTAGCTCCTTCGGTCTCCTCGATCGTTGTGAGAAGTA
AGTTGGCCGAGTGTTATCCTCATGTTTATGGCAGCACTGCATAATTCTCTTACTGTCTATG
CCATCCGTAAGATGCTTTTTCTGTGACTGGTGAGTACTCAACCAAGTCATTCTGAGAATAGTG
TATGCGGCGACCGAGTTGCTCTTGCCCGGCGTCAATACGGGATAATACCGCGCCACATAGCA
GAACTTTAAAAGTGCTCATCATTGGAACCGTTCTTCGGGGCGAAAACCTCTCAAGGATCTTA
CCGCTGTTGAGATCCAGTTTCGATGTAACCCACTCGTGACCCCAACTGATCTTCAGCATCTTT
TACTTTTACCAGCGTTTCTGGGTGAGCAAAAACAGGAAGGCAAAATGCCGCAAAAAAGGGAA
TAAGGGCGACACGGAAATGTTGAATACTCATACTCTTCCTTTTTCAATATTATTGAAGCATT
TATCAGGGTTATTGTCTCATGAGCGGATACATATTTGAATGTATTTAGAAAAATAACAAAT
AGGGGTTCCGCGCACATTTCCCGAAAAAGTGCCACCTGACGCGCCCTGTAGCGGCGCATTAA
GCGCGGCGGGTGTGGTGGTTACGCGCAGCGTGACCGCTACACTTGCCAGCGCCCTAGCGGCC
GCTCCTTTTCGTTTTCTTCCCTTCCCTTCTCGCCACGTTCCGCGGCTTTCCCGTCAAGCTCT
AAATCGGGGGCTCCCTTTAGGGTTCCGATTTAGTGCTTTACGGCACCTCGACCCCCAAAAAC
TTGATTAGGGTGATGGTTACGTTAGTGGGCCATCGCCCTGATAGACGGTTTTTCGCCCTTTG
ACGTTGGAGTCCACGTTCTTTAATAGTGGACTCTTGTTCCAACTGGAACAACACTCAACCC
TATCTCGGTCTATTCTTTTGATTTTATAAGGGATTTTGCCGATTTTCGGCCTATTGGTTAAAAA
ATGAGCTGATTTAACAAAAATTTAACGCGAATTTTAACAAAAATATTAACGTTTACAATTTCC
CATTGCGCATTCAGGCTGCGCAACTGTTGGGAAGGGCGATCGGTGCGGGCCTCTTCGCTATT
ACGCCAGCCCAAGCTACCATGATAAGTAAGTAATATTAAGGTACGGGAGGTACTTGGAGCGG
CCGCAATAAAATATCTTTATTTTATTACATCTGTGTGTTGGTTTTTTGTGTGAATCGATAG
TACTAACATACGCTCTCCATCAAAACAAAACGAAACAAAACAACTAGCAAAATAGGCTGTC
CCCAGTGCAAGTGACAGGTGCCAGAACATTT

Fig. 10

pCMV-luc-attP (SEQ ID NO: 3)

CTCTATCGATAGGTACCGAGCTCTTACGCGTGCTAGCCCTCGAGCAGGATCTATACATTGAA
TCAATATTGGCAATTAGCCATATTAGTCATTGGTTATATAGCATAAATCAATATTGGCTATT
GGCCATTGCATACGTTGTATCTATATCATAATATGTACATTTATATTGGCTCATGTCCAATA
TGACCGCCATGTTGACATTGATTATTGACTAGTTATTAATAGTAATCAATTACGGGGTCATT
AGTTCATAGCCCATATATGGAGTTCGCGTTACATAACTTACGGTAAATGGCCCGCCTGGCT
GACCGCCCAACGACCCCCGCCCATTGACGTCAATAATGACGTATGTTCCCATAGTAACGCCA
ATAGGGACTTTCCATTGACGTCAATGGGTGGAGTATTTACGGTAAACTGCCCACTTGGCAGT
ACATCAAGTGTATCATATGCCAAGTCCGCCCCCTATTGACGTCAATGACGGTAAATGGCCCG
CCTGGCATTATGCCCAGTACATGACCTTACGGGACTTTCCCTACTTGGCAGTACATCTACGTA
TTAGTCATCGCTATTACCATGGTGATGCGGTTTTTGGCAGTACATCAATGGGCGTGGATAGCG
GTTTGAACACGAGGATTTCCTCAAGTCTCCACCCCATTGACGTCAATGGGAGTTTGTTCCTGGC
ACCAAAATCAACGGGACTTTCCAAAATGTCGTAACAACTCCGCCCCATTGACGCAAAATGGGC
GGTAGGCGTGTACGGTGGGAGGTCTATATAAGCAGAGCTCGTTAGTGAACCGTCAGATCGC
CTGGAGACGCCATCCACGCTGTTTTGACCTCCATAGAAGACACCGGGACCGATCCAGCCTCC
CCTCGAAGCTCGACTCTAGGGGCTCGAGATCTGCGATCTAAGTAAGCTTGGCATTCCGGTAC
TGTTGGTAAAGCCACCATGGAAGACGCCAAAACATAAAGAAAGGCCCGCGCCATTCTATC
CGCTGGAAGATGGAACCGCTGGAGAGCAACTGCATAAGGCTATGAAGAGATACGCCCTGGTT
CTGGAACAATTGCTTTTACAGATGCACATATCGAGGTGGACATCACTTACGCTGAGTACTT
CGAAATGTCCGTTTCGGTTGGCAGAAGCTATGAAACGATATGGGCTGAATACAAATCACAGAA
TCGTCGTATGCAGTGAAGAACTCTCTTCAATTCTTTATGCGGTTGTTGGGCGCTTATTTATC
GGAGTTGCAGTTGCGCCCGGAACGACATTTATAATGAACGTGAATTGCTCAACAGTATGGG
CATTTTCGCAGCCTACCGTGGTGTTCGTTTTCCAAAAGGGGTTGCAAAAAATTTTGAACGTGC
AAAAAAGCTCCCAATCATCCAAAAAATTATTATCATGGATTCTAAAACGGATTACCAGGGA
TTTCAGTCGATGTACACGTTTCGTACATCTCATCTACCTCCCGGTTTTTAATGAATACGATTT
TGTGCCAGAGTCCCTCGATAGGGACAAGACAATTGCACTGATCATGAACTCCTCTGGATCTA
CTGGTCTGCCTAAAGGTGTGCTCTGCCTCATAGAACTGCCTGCGTGAGATTCTCGCATGCC
AGAGATCCTATTTTTTGGCAATCAAATCATTTCCGGATACTGCGATTTTAAAGTGTGTTCATT
CCATCACGGTTTTTGAATGTTTACTACACTCGGATATTTGATATGTGGATTTTCGAGTCGTCT
TAATGTATAGATTTGAAGAAGAGCTGTTTCTGAGGAGCCTTCAGGATTACAAGATTCAAAGT
GCGCTGCTGGTGCCAACCCTATTCTCCTTCTTCGCCAAAAGCACTCTGATTGACAAATACGA
TTTATCTAATTTACACGAAATTGCTTCTGGTGGCGCTCCCCCTCTCTAAGGAAGTCGGGGAAG
CGTTTGCCAAGAGGTTCCATCTGCCAGGTATCAGGCAAGGATATGGGCTCACTGAGACTACA
TCAGCTATTCTGATTACACCCGAGGGGATGATAAACCGGGCGCGGTGCGTAAAGTTGTTCC
ATTTTTTGAAGCGAAGGTTGTGGATCTGGATACCGGGAAAACGCTGGGCGTTAATCAAAGAG
GCGAAGTGTGTGTGAGAGGTCTATGATTATGTCCGGTTATGTAAACAATCCGGAAGCGACC
AACGCCTTGATTGACAAGGATGGATGGCTACATTCTGGAGACATAGCTTACTGGGACGAAGA
CGAACACTTCTTCATCGTTGACCGCCTGAAGTCTCTGATTAAGTACAAAGGCTATCAGGTGG
CTCCCGCTGAATTGGAATCCATCTTGCTCCAACACCCCAACATCTTCGACGCAGGTGTCGCA
GGTCTTCCCGACGATGACGCGCGTGAACCTCCCGCCGCGTTGTTGTTTTGGAGCACGGAAA
GACGATGACGGAAAAAGAGATCGTGGATTACGTGCGCAGTCAAGTAACAACCGCGAAAAAGT
TGCGCGGAGGAGTTGTGTTTTGTGGACGAAGTACCGAAAGGTCCTTACCGGAAAACCTCGACGCA
AGAAAAATCAGAGAGATCCTCATAAAGGCCAAGAAGGGCGGAAAGATCGCCGTGTAATTCTA
GAGTCGGGGCGGCCGCGCTTCGAGCAGACATGATAAGATACATTGATGAGTTTGGACAAA
CCACAAC TAGAATGCAGTGAAAAAATGCTTTATTTGTGAAATTTGTGATGCTATTGCTTTA
TTTGTAAACATTATAAGCTGCAATAAACAAGTTAAACAACAATTCGATTCAATTTTATGTT
TCAGGTTTCAGGGGGAGGTGTGGGAGGTTTTTTAAAGCAAGTAAACCTCTACAAATGTGGTA
AAATCGATAAGGATCAATTCGGCTTCGACTAGTACTGACGGACACACCGAAGCCCCGGCGGC
AACCTCAGCGGATGCCCCGGGGCTTCACGTTTTCCAGGTCAGAAGCGGTTTTTCGGGAGTA
GTGCCCCAAGTGGGGTAACCTTTGAGTTCTCTCAGTTGGGGGCGTAGGGTCGCCGACATGAC
ACAAGGGGTTGTGACCGGGGTGGACACGTACGCGGGTGCTTACGACCGTCAGTCGCGCGAGC
GCGACTAGTACAAGCCGAATTGATCCGTCGACCGATGCCCTTGAGAGCCTTCAACCCAGTCA
GCTCCTTCCGGTGGGCGCGGGGCATGACTATCGTCGCCGCACTTATGACTGTCTTCTTTATC
ATGCAACTCGTAGGACAGGTGCCGCGAGCGCTCTCCGCTTCTCGCTCACTGACTCGCTGC

GCTCGGTCGTTTCGGCTGCGGCGAGCGGTATCAGCTCACTCAAAGGCGGTAAATACGGTTATCC
ACAGAATCAGGGGATAACGCAGGAAAGAACATGTGAGCAAAAGGCCAGCAAAAGGCCAGGAA
CCGTAAAAAGGCCGCGTTGCTGGCGTTTTTCCATAGGCTCCGCCCCCTGACGAGCATCACA
AAAATCGACGCTCAAGTCAGAGGTGGCGAAACCCGACAGGACTATAAAGATACCAGGCGTTT
CCCCCTGGAAGCTCCCTCGTGCGCTCTCTGTTCCGACCCCTGCCGCTTACCGGATACCTGTC
CGCCTTTCTCCCTTCGGGAAGCGTGGCGCTTTCTCAATGCTCACGCTGTAGGTATCTCAGTT
CGGTGTAGGTGCTTCGCTCCAAGCTGGGCTGTGTGCACGAACCCCCCGTTACGCCCCGACCGC
TGCGCCTTATCCGGTAACTATCGTCTTGAGTCCAACCCGGTAAGACACGACTTATCGCCACT
GGCAGCAGCCACTGGTAACAGGATTAGCAGAGCGAGGTATGTAGGCGGTGCTACAGAGTTCT
TGAAGTGGTGGCTTAACCTACGGCTACACTAGAAGGACAGTATTTGGTATCTGCGCTCTGCTG
AAGCCAGTTACCTTCGGAAAAAGAGTTGGTAGCTCTTGATCCGGCAAACAAACCCGCTGG
TAGCGGTGGTTTTTTTTGTTTGCAAGCAGCAGATTACGCGCAGAAAAAAGGATCTCAAGAAG
ATCCTTTGATCTTTTTCTACGGGTCTGACGCTCAGTGGAAACGAAACTCACGTTAAGGGATT
TTGGTCATGAGATTATCAAAAAGGATCTTCACCTAGATCCTTTTTAAATTAAAAATGAAGTTT
TAAATCAATCTAAAGTATATATGAGTAAACTTGGTCTGACAGTTACCAATGCTTAATCAGTG
AGGCACCTATCTCAGCGATCTGTCTATTTTCGTTTCATCCATAGTTGCCTGACTCCCCGTCGTG
TAGATAACTACGATACGGGAGGGCTTACCATCTGGCCCCAGTGCTGCAATGATACCGCGAGA
TCCACGCTTACCGGCTCCAGATTATCAGCAATAAACAGCCAGCCGGAAGGCCGAGCGCA
GAAGTGGTCTTGCAACTTTATCCGCCTCCATCCAGTCTATTAATTGTTGCCGGAAGCTAGA
GTAAGTAGTTTCGCCAGTTAATAGTTTGCACAACGTTGTTGCCATTGCTACAGGCATCGTGGT
GTCACGCTCGTCGTTTGGTATGGCTTCATTCAGCTCCGGTTCCTCAACGATCAAGGCGAGTTA
CATGATCCCCCATGTTGTGCAAAAAGCGGTTAGCTCCTTCGGTCTCCTCGATCGTTGTCAGA
AGTAAGTTGGCCGCGAGTGTTATCACTCATGGTTATGGCAGCACTGCATAATTCTCTTACTGT
CATGCCATCCGTAAGATGCTTTTCTGTGACTGGTGAGTACTCAACCAAGTCATTCTGAGAAT
AGTGATGCGGCGACCGAGTTGCTCTTGGCCGCGCTCAATACGGGATAATACCGCGCCACAT
AGCAGAACTTTAAAAGTGCTCATCATTGGAAAACGTTCTTCGGGGCGAAAACCTCTCAAGGAT
CTTACCGCTGTTGAGATCCAGTTCGATGTAACCCACTCGTGCACCCAACTGATCTTCAGCAT
CTTTTACTTTTACCAGCGTTTCTGGGTGAGCAAAAACAGGAAGGCAAAATGCCGCAAAAAAG
GGAATAAGGGCGACACGGAAATGTTGAATACTCATACTCTTCTTTTCAATATTATTGAAG
CATTTATCAGGGTTATTGTCTCATGAGCGGATACATATTTGAATGTATTTAGAAAAATAAAC
AAATAGGGGTTCCGCGCACATTTCCCCGAAAAGTGCCACCTGACGCGCCCTGTAGCGGCGCA
TTAAGCGCGGCGGTGTGGTGGTTACGCGCAGCGTGACCGCTACACTTGCCAGCGCCCTAGC
GCCCCGCTCCTTTCGCTTTCTTCCCTTCTTCTCGCCACGTTCCGCGGCTTTCCCCGTCAAG
CTCTAAATCGGGGGCTCCCTTTAGGGTTCCGATTTAGTGCTTTACGGCACCTCGACCCCCAA
AACTTGATTAGGGTGATGGTTACGTTAGTGGGCCATCGCCCTGATAGACGGTTTTTCGCCC
TTTGACGTTGGAGTCCACGTTCTTTAATAGTGGACTCTTGTTCCAAACTGGAACAACACTCA
ACCCTATCTCGGTCTATTCTTTTGATTTATAAGGGATTTTGCCGATTTCCGGCCTATTGGTTA
AAAAATGAGCTGATTTAACAAAAATTTAACGCGAATTTAACAAAAATATTACGTTTACAAT
TTCCCATTCGCCATTTCAGGCTGCGCAACTGTTGGGAAGGGCGATCGGTGCGGGCCTCTTCGC
TATTACGCCAGCCCAAGCTACCATGATAAGTAAGTAATATTAAAGGTACGGGAGGTACTTTGGA
GCGGCCGCAATAAAATATCTTTATTTTATTACATCTGTGTGTTGGTTTTTTGTGTGAATCG
ATAGTACTAACATACGCTCTCCATCAAAACAAAACGAAACAAAACAACTAGCAAAATAGGC
TGTCGCCAGTGCAAGTGCCAGGTGCCAGAACATTT

Fig. 11

pCMV-pur-attB (SEQ ID NO: 4)

CTAGAGTCGGGGCGGCCGGCCGCTTCGAGCAGACATGATAAGATACATTGATGAGTTTGGAC
AAACCACAAC TAGAATGCAGTGAAAAAATGCTTTATTTGTGAAATTTGTGATGCTATTGCT
TTATTTGTAACCATTATAAGCTGCAATAACAAGTTAACAACAACAATTGCATTCATTTTAT
GTTTCAGGTTTCAGGGGGAGGTGTGGGAGGTTTTTTAAAGCAAGTAAAACCTCTACAAATGTG
GTAAAAATCGATAAGGATCAATTCGGCTTCAGGTACCGTCGACGATGTAGGTACGGTCTCGA
AGCCGCGGTGCGGGTGCCAGGGCGTGCCCTTGGGCTCCCCGGGCGCGTACTCCACCTCACCC
ATCTGGTCCATCATGATGAACGGGTTCGAGGTGGCGGTAGTTGATCCC GGCGAACGCGCGGCG
CACCGGGAAGCCCTCGCCCTCGAAACCGCTGGGCGCGGTGGTCACGGTGAGCACGGGACGTG
CGACGGCGTTCGGCGGGTGCGGATACGCGGGGCGAGCTCAGCGGGTTCTCGACGGTCACGGCG
GGCATGTTCGACAGCCGAATTGATCCGTTCGACCGATGCCCTTGAGAGCCTTCAACCCAGTCAG
TCCCTTCGGGTGGGCGCGGGGCATGACTATCGTCGCGCACTTATGACTGTCTTCTTTATCA
TGCAACTCGTAGGACAGGTGCCGGCAGCGCTCTTCCGCTTCTTCGCTCACTGACTCGCTGCG
CTCGGTCTGTTTCGGCTGCGGCGAGCGGTATCAGCTCACTCAAAGGCGGTAATACGGTTATCCA
CAGAATCAGGGGATAACGCAGGAAAGAACATGTGAGCAAAAGGCCAGCAAAAGGCCAGGAAC
CGTAAAAAGGCCGCGTTGCTGGCGTTTTTCCATAGGCTCCGCCCCCTGACGAGCATCACAA
AAATCGACGCTCAAGTCAGAGGTGGCGAAACCCGACAGGACTATAAAGATACCAGGCGTTTC
CCCCTGGAAGCTCCCTCGTGCGCTCTCTGTTCGACCCCTGCCGCTTACCGGATACCTGTCC
GCCCTTCTCCCTTCGGGAAGCGTGGCGCTTTCTCAATGCTCACGCTGTAGGTATCTCAGTTT
GGTGTAGGTCTGCTCCAAGCTGGGCTGTGTGCACGAACCCCCGTTTCAGCCCGACCGCT
GCGCCTTATCCGTAACCTATCGTCTTGAGTCCAACCCGGTAAGACACGACTTATCGCCACTG
GCAGCAGCCACTGGTAACAGGATTAGCAGAGCGAGGTATGTAGGCGGTGCTACAGAGTTCTT
GAAGTGGTGGCCTAACTACGGCTACACTAGAAGGACAGTATTTGGTATCTGCGCTCTGCTGA
AGCCAGTTACCTTCGGAAAAAGAGTTGGTAGCTCTTGATCCGGCAAACAACACCCGCTGGT
AGCGGTGGTTTTTTTTGTTTGCAAGCAGCAGATTACGCGCAGAAAAAAGGATCTCAAGAAGA
TCCTTTGATCTTTTTCTACGGGGTCTGACGCTCAGTGGAACGAAACTCACGTTAAGGGATTT
TGGTCATGAGATTATCAAAAAGGATCTTCACCTAGATCCTTTTAAATTAATAAATGAAGTTTT
AAATCAATCTAAAGTATATATGAGTAACTTGGTCTGACAGTTACCAATGCTTAATCAGTGA
GGCACCTATCTCAGCGATCTGTCTATTTCTGTTTCATCCATAGTTGCCTGACTCCCCGTCGTGT
AGATAACTACGATACGGGAGGGCTTACCATCTGGCCCCAGTGCTGCAATGATACCGCGAGAC
CCACGCTCACCGGCTCCAGATTTATCAGCAATAAACACAGCCAGCCGGAAGGGCCGAGCGCAG
AAGTGGTCCCTGCAACTTATCCGCCTCCATCCAGTCTATTAAATTGTTGCCGGAAGCTAGAG
TAAGTAGTTTCGCCAGTTAATAGTTTGGCGCAACGTTGTTGCCATTGCTACAGGCATCGTGGTG
TCACGCTCGTCTGTTTGGTATGGCTTCATTCAGCTCCGGTTCCTAACGATCAAGGCGAGTTAC
ATGATCCCCCATGTTGTGCAAAAAAGCGTTAGCTCCTTCGGTCTCTCCGATCGTTGTGAGAA
GTAAGTTGGCCGAGTGTTATCACTCATGGTTATGGCAGCACTGCATAATTCTCTTACTGTC
ATGCCATCCGTAAGATGCTTTTCTGTGACTGGTGAAGTCAACCAAGTCATTCTGAGAATA
GTGTATGCGGCGACCGAGTTGCTCTTGCCCGGCGTCAATACGGGATAATACCGCGCCACATA
GCGAACTTTTAAAGTGCTCATCATTGGAAAAACGTTCTTTCGGGGCGAAAACTCTCAAGGATC
TTACCGCTGTTGAGATCCAGTTCGATGTAACCCACTCGTGACCCCACTGATCTTCAGCATC
TTTTACTTTTACCAGCGTTTCTGGGTGAGCAAAAAACAGGAAGGCAAAATGCCGCAAAAAAGG
GAATAAGGGCGACACGGAAATGTTGAATACTCATACTCTTCTTTTCAATATTATTGAAGC
ATTTATCAGGGTTATTGTCTCATGAGCGGATACATATTTGAATGTATTTAGAAAAATAACA
AATAGGGGTTCCGCGCACATTTCCCCGAAAAGTGCCACCTGACGCGCCCTGTAGCGGCGCAT
TAAGCGCGGCGGTGTGGTGGTTACGCGCAGCGTGACCGCTACACTTGCCAGCGCCCTAGCG
CCGCTCCTTTTCGCTTCTTCCCTTCTTCTCGCCACGTTTCGCGGCTTTCCCCGTCAAGC
TCTAAATCGGGGGCTCCC'TTTAGGGTTCCGATTTAGTGCTTTACGGCACCTCGACCCCAAAA
AACTTGATTAGGGTGATGGTTCACGTAGTGGCCATCGCCCTGATAGACGGTTTTTTCGCCCT
TTGACGTTGGAGTCCACGTTCTTTAATAGTGGACTCTTGTTCCTAACTGGAACAACACTCAA
CCCTATCTCGGTCTATTCTTTTGATTTATAAGGGATTTTGCCGATTTCCGGCTATTGGTTAA
AAAATGAGCTGATTTAAACAAAAATTTAACGCGAATTTTAAACAAAATATTAACGTTTACAATT
TCCCATTCGCCATTCAGGCTGCGCAACTGTTGGGAAGGGCGATCGGTGCGGGGCTTTCGCT
ATTACGCCAGCTTAAGCTACCATGATAAGTAAGTAATTAAGGTACGGGAGTACTTGGAG
CGGCCGCAATAAAATATCTTTATTTTTCATTACATCTGTGTGTTGGTTTTTTGTGTGAATCGA

TAGTACTAACATACGCTCTCCATCAAAACAAAACGAAACAAAACAAACTAGCAAAATAGGCT
GTCCCCAGTGCAAGTGCAGGTGCCAGAACATTTCTCTATCGATAGGTACCGAGCTCTTACGC
TTGGTTATATAGCATAAATCAATATTGGCTATTGGCCATTGCATACGTTGTATCTATATCAT
AATATGTACATTTATATTGGCTCATGTCCAATATGACCGCCATGTTGACATTGATTATTGAC
TAGTTATTAATAGTAAATCAATTACGGGGTCATTAGTTTCATAGCCCATATATGGAGTTCGCGC
TTACATAACTTACGGTAAATGGCCCGCCTGGCTGACCGCCCAACGACCCCCCGCCATTGACG
TCAATAATGACGTATGTTCCCATAGTAACGCCAATAGGGACTTTCCATTGACGTCAATGGGT
GGAGTATTTACGGTAAACTGCCCCACTTGGCAGTACATCAAGTGTATCATATGCCAAGTCCGC
CCCCTATTGACGTCAATGACGGTAAATGGCCCGCCTGGCATTATGCCCAGTACATGACCTTA
CGGGACTTTTCTACTTGGCAGTACATCTACGTATTAGTCATCGCTATTACCATGGTGTATGCG
GTTTTGGCAGTACATCAATGGGCGTGGATAGCGGTTTGACTCACGGGGATTTCCAAGTCTCC
ACCCCATTTGACGTCAATGGGAGTTTGTGTTTGGCACCAAAATCAACGGGACTTTCCAAAATGT
CGTAACAACCTCCGCCCCATTGACGCAAATGGGCGGTAGGCGTGTACGGTGGGAGGTCTATAT
AAGCAGAGCTCGTTTAGTGAACCGTCAGATCGCCTGGAGACGCCATCCACGCTGTTTTGACC
TCCATAGAAGACACCGGGACCGATCCAGCCTCCCCTCGAAGCTCGACTCTAGGGGGCTCGAGA
TCTGCGATCTAAGTAAGCTTGATGCCTGCAGGTGCGCCGCCACGACCGGTGCCGCCACCAT
CCCCTGACCCACGCCCCCTGACCCCTCACAAGGAGACGACCTTCCATGACCGAGTACAAGCCC
ACGGTGCGCCTCGCCACCCGCGACGACGTCCCCCGGGCGGTACGCACCCTCGCCGCCGCGTT
CGCCGACTACCCCGCCACGCGCCACACCGTCGACCCGGACCGCCACATCGAGCGGGTCACCG
AGCTGCAAGAACTCTTCCTCACGCGCGTCGGGCTCGACATCGGCAAGGTGTGGGTGCGCGAC
GACGGCGCCGCGGTGGCGGTCTGGACCACGCCGGAGAGCGTCGAAGCGGGGGCGGTGTTTCGC
CGAGATCGGCCCCGCGCATGGCCGAGTTGAGCGGTTCCCGGCTGGCCGCGCAGCAACAGATGG
AAGGCCTCCTGGCGCCGCACCGGCCCAAGGAGCCCGCGTGGTTCTTGGCCACCGTCGGCGTCT
TCGCCCCGACCACGAGGGCAAGGTTCTGGGCAGCGCCGTCGTGCTCCCCGGAGTGGAGGCGGC
CGAGCGCGCCGGGGTGCCCGCCTTCTTGAGACCTCCGCGCCCCGCAACCTCCCCTTCTACG
AGCGGCTCGGCTTCACCGTCACCGCCGACGTGAGGTGCCCGAAGGACCGCGCACCTGGTGC
ATGACCCGCAAGCCCGGTGCCTGACGCCCCGCCACGACCCGCGAGCGCCGACCGAAAGGAG
CGCACGACCCCATGGCTCCGACCGAAGCCGACCCGGGCGGCCCCGCGGACCCCGCACCCGCC
CCCCGAGGCCACCGACT

Fig. 12

pCMV-pur-attP (SEQ ID NO: 5)

CTAGAGTCGGGGCGGCCGGCCGCTTCGAGCAGACATGATAAGATACATTGATGAGTTTGGAC
AAACCACAACCTAGAATGCAGTGAAAAAATGCTTTATTTGTGAAATTTGTGATGCTATTGCT
TTATTTGTAACCATTATAAGCTGCAATAAACAAGTTAACAACAACAATTGCATTCATTTTAT
GTTTCAGGTTTCAAGGGGAGGTGTGGGAGGTTTTTTAAAGCAAGTAAAACCTCTACAAATGTG
GTAAAATCGATAAGGATCAATTCCGGCTTCGACTAGTACTGACGGACACACCGAAGCCCCGGC
GGCAACCCTCAGCGGATGCCCCGGGGCTTCACGTTTTCCAGGTGAGAAGCGGTTTTTCGGGA
GTAGTGCCCCAAGTGGGGTAACCTTTGAGTTCTCTCAGTTGGGGGCGTAGGGTCGCCGACAT
GACACAAGGGGTTGTGACCGGGGTGGACACGTACGCGGGTGCTTACGACCGTCAGTCGCGCG
AGCGCGACTAGTACAAGCCGAATTGATCCGTCGACCGATGCCCTTGAGAGCCTTCAACCCAG
TCAGCTCCTTCCGGTGGGCGCGGGGCATGACTATCGTCGCCGCACTTATGACTGTCTTCTTT
ATCATGCAACTCGTAGGACAGGTGCGGCGAGCGCTCTTCCGCTTCTCGCTCACTGACTCGC
TGCGCTCGGTCTGTTCCGGCTGCGGCGAGCGGTATCAGCTCACTCAAAGGCGGTAAATACGGTTA
TCCACAGAATCAGGGGATAACGCAGGAAAGAACATGTGAGCAAAGGCCAGCAAAGGCCAG
GAACCGTAAAAAGGCCCGCTTGCTGGCGTTTTTCCATAGGCTCCGCCCCCTGACGAGCATC
ACAAAAATCGACGCTCAAGTCAGAGGTGGCGAAACCCGACAGGACTATAAAGATACCAGGCG
TTTCCCCCTGGAAGCTCCCTCGTGCGCTCTCCTGTTCCGACCCTGCCGCTTACCGGATACCT
GTCCGCCTTTCTCCCTTCGGGAAGCGTGGCGCTTCTCAATGCTCACGCTGTAGGTATCTCA
GTTTCGGTCATGAGATTATCAAAAAGGATCTTCACCTAGATCCTTTTAAATTTAAAAATGAAG
CGCTGCGCCTTATCCGGTAACCTATCGTCTTGAGTCCAACCCGGTAAGACACGACTTATCGCC
ACTGGCAGCAGCCACTGGTAACAGGATTAGCAGAGCGAGGTATGTAGGCGGTGCTACAGAGT
TCTTGAAGTGGTGGCCTAACTACGGCTACACTAGAAGGACAGTATTTGGTATCTGCGCTCTG
CTGAAGCCAGTTACCTTCGGAAAAAGAGTTGGTAGCTCTTGATCCGGCAAACAAACCACCGC
TGGTAGCGGTGGTTTTTTTTGTGTTGCAAGCAGCAGATTACGCGCAGAAAAAAGGATCTCAAG
AAGATCCTTTGATCTTTTCTACGGGGTCTGACGCTCAGTGGAACGAAAACCTCACGTTAAGGG
ATTTTGGTCATGAGATTATCAAAAAGGATCTTCACCTAGATCCTTTTAAATTTAAAAATGAAG
TTTTAAATCAATCTAAAGTATATATGAGTAACTTGGTCTGACAGTTACCAATGCTTAATCA
GTGAGGCACCTATCTCAGCGATCTGTCTATTTCTGTTTATCCATAGTTGCCCTGACTCCCCGTC
GTGTAGATAACTACGATACGGGAGGGCTTACCATCTGGCCCCAGTGCTGCAATGATACCGCG
AGACCCACGCTCACCAGGCTCCAGATTTATCAGCAATAAACCAGCCAGCCGGAAGGGCCGAGC
GCAGAAGTGGTCTGCAACTTTATCCGCCTCCATCCAGTCTATTAATTGTTGCCGGGAAGCT
AGAGTAAGTAGTTTCGCCAGTTAATAGTTTGCACAACGTTGTTGCCATTGCTACAGGCATCGT
GGTGTACGCTCGTCTGTTTGGTATGGCTTCATTACGCTCCGGTTCCCAACGATCAAGGCGAG
TTACATGATCCCCCATGTTGTGCAAAAAGCGGTTAGCTCCTTCCGGTCTCCGATCGTTGTC
AGAAGTAAGTTGGCCGAGTGTTATCACTCATGGTTATGGCAGCACTGCATAATTCTCTTAC
TGTCATGCCATCCGTAAGATGCTTTTCTGTGACTGGTGAGTACTCAACCAAGTCATTCTGAG
AATAGTGTATGCGGCGACCCGAGTTGCTCTTGCCCGGCGTCAATACGGGATAATACCGCGCCA
CATAGCAGAACTTTAAAGTGCTCATCATTGAAAAACGTTCTTCGGGGCGAAAACCTCTCAAG
GATCTTACCCTGTTGAGATCCAGTTCGATGTAACCCACTCGTGACCCCACTGATCTTCAG
CATCTTTTACTTTTACCAGCGTTTCTGGGTGAGCAAAAACAGGAAGGCAAAATGCCGCAAAA
AAGGGAATAAGGGCGACACGGAAATGTTGAATACTCATACTCTTCCTTTTTCAATATTATTG
AAGCATTTATCAGGGTTATTGTCTCATGAGCGGATACATATTTGAATGTATTTAGAAAAATA
AACAAATAGGGGTTCCGCGCACATTTCCCCGAAAAGTGCCACCTGACGCGCCCTGTAGCGGC
GCATTAAGCGCGGCGGGTGTGGTGGTTACGCGCAGCGTGACCGCTACACTTGCCAGCGCCCT
AGCGCCCCGCTCCTTTCGCTTCTTCCCTTCTTCTCGCCACGTTCCGCCGCTTTCCTCCGTC
AAGCTCTAAATCGGGGGCTCCCTTTAGGGTTCCGATTTAGTGCTTTACGGCACCTCGACCCC
AAAAAAGTTGATTAGGGGTGATGGTTACAGTAGTGGGCCATCGCCCTGATAGACGGTTTTTCG
CCCTTTGACGTTGGAGTCCACGTTCTTTAATAGTGGACTCTTGTTCCAACTGGAACAACAC
TCAACCCTATCTCGGTCTATTCTTTTGATTTATAAGGGATTTTGCCGATTTCCGGCCTATTGG
TTAAAAAATGAGCTGATTTAACAATAATTTAACGCGAATTTTAAACAAAATATTAACGTTTAC
AATTTCCCATTCGCCATTACAGGCTGCGCAACTGTTGGGAAGGGCGATCGGTGCGGGCCTCTT
CGCTATTACGCCAGCCCAAGCTACCATGATAAGTAAGTAATATTAAGGTACGGGAGGTACTT
GGAGCGGCCGCAATAAAATATCTTTATTTTCAATACATCTGTGTGTTGTTGTTGTTGTA
TCGATAGTACTAACATACGCTCTCCATCAAAACAAAACGAAACAAAACAACTAGCAAAATA

GGCTGTCCCCAGTGCAAGTGCAGGTGCCAGAACATTTCTCTATCGATAGGTACCGAGCTCTT
ACGCGTGTAGCCCTCGAGCAGGATCTATACATTGAATCAATATTGGCAATTAGCCATATTA
GTCATTGGTTATATAGCATAAATCAATATTGGCTATTGGCCATTGCATACGTTGTATCTATA
TCATAATATGTACATTTATATTGGCTCATGTCCAATATGACCGCCATGTTGACATTGATTAT
TGACTAGTTATTAATAGTAATCAATTACGGGGTCATTAGTTCATAGCCCATATATGGAGTTC
CGCGTTACATAAATTACGGTAAATGGCCCGCCTGGCTGACCGCCCAACGACCCCGCCCAT
GACGTCAATAATGACGTATGTTCCCATAGTAACGCCAATAGGGACTTTCCATTGACGTCAAT
GGGTGGAGTATTTACGGTAACTGCCCACTTGGCAGTACATCAAGTGTATCATATGCCAAGT
CCGCCCCCTATTGACGTCAATGACGGTAAATGGCCCGCCTGGCATTATGCCCAGTACATGAC
CTTACGGGACTTTTCTACTTGGCAGTACATCTACGTATTAGTCATCGCTATTACCATGGTGA
TGCGGTTTTTGGCAGTACATCAATGGGCGTGGATAGCGGTTTGACTCACGGGGATTTCCAAGT
CTCCACCCCATTTGACGTCAATGGGAGTTTGTTTTGGCACCAAAATCAACGGGACTTTCCAAA
ATGTCGTAACAACCTCCGCCCCATTGACGCAAATGGGCGGTAGGCGTGTACGGTGGGAGGTCT
ATATAAGCAGAGCTCGTTTTAGTGAACCGTCAGATCGCCTGGAGACGCCATCCACGCTGTTTT
GACCTCCATAGAAGACACCGGGACCGATCCAGCCTCCCCTCGAAGCTCGACTCTAGGGGCTC
GAGATCTGCGATCTAAGTAAGCTTGTCATGCCTGCAGGTGCGCCGCCACGACCGGTGCCGCCA
CCATCCCCTGACCCACGCCCCTGACCCCTCACAAGGAGACGACCTTCCATGACCGAGTACAA
GCCCCGCGTGCCTCGCCACCCGCGACGACGTCCCCGGGCGGTACGCACCCTCGCCGCCG
CGTTGCGCGACTACCCCGCCACGCGCCACACCGTCGACCCGGACCGCCACATCGAGCGGGTC
ACCGAGCTGCAAGAACTCTTCTCACGCGCGTCTGGGCTCGACATCGGCAAGGTGTGGGTGCG
GGACGACGGCGCCGCGGTGGCGGTCTGGACCACGCCGGAGAGCGTCGAAGCGGGGGCGGTGT
TCGCCGAGATCGGCCCGCGCATGGCCGAGTTGAGCGGTTCCCGGCTGGCCGCGCAGCAACAG
ATGGAAGGCCTCCTGGCGCCGCACCGGCCAAGGAGCCCGCGTGTTCTCGCCACCGTCGG
CGTCTCGCCGACCACCAGGGCAAGGGTCTGGGCAGCGCCGTGCTGCTCCCCGGAGTGAGG
CGCCGAGCGCGCCGGGTGCCCGCCTTCTGGAGACCTCCGCGCCCCGCAACCTCCCCTTC
TACGAGCGGCTCGGCTTCACCGTCACCGCCGACGTGAGGTGCCCCGAAGGACCGCGCACCTG
GTGCATGACCCGCAAGCCCGGTGCCTGACGCCCCGCCACGACCCGCGAGCGCCGACCGAAA
GGAGCGCACGACCCCATGGCTCCGACCGAAGCCGACCCGGCGGCCCGCGACCCCGCACCC
CGCCCCCGAGGCCACCGACT

Fig. 13

pCMV-EGFP-attB (SEQ ID NO: 6)

CTAGAGTCGGGGCGGCCGGCCGCTTCGAGCAGACATGATAAGATACATTGATGAGTTTGGAC
AAACCACAACCTAGAAATGCAGTGAAAAAATGCTTTATTTGTGAAATTTGTGATGCTATTGCT
TTATTTGTAACCATTTATAAGCTGCAATAAACAAGTTAACAACAACAATTGCATTCAATTTTAT
GTTTCAGGTTTCAGGGGGAGGTGTGGGAGGTTTTTTAAAGCAAGTAAACCTCTACAAATGTG
GTAAATCGATAAGGATCAATTCGGCTTCAGGTACCGTCGACGATGTAGGTCACGGTCTCGA
AGCCGCGGTGCGGGTGCCAGGGCGTGCCCTTGGGCTCCCCGGGCGCGTACTCCACCTCACCC
ATCTGGTCCATCATGATGAACGGGTGCGAGGTGCGGTAAGTTGATCCCGCGAACGCGCGGGC
CACCGGGAAGCCCTCGCCCTCGAAACCGCTGGGCGCGGTGGTCACGGTGAGCACGGGACGTG
CGACGGCGTTCGGCGGGTGCGGATACGCGGGGCGAGCGTCAGCGGGTTCTCGACGGTCACGGCG
GGCATGTGCGACAGCCGAATTGATCCGTGCGACCGATGCCCTTGAGAGCCTTCAACCCAGTCAG
CTCCTTCCGGTGGGCGCGGGGCATGACTATCGTCGCCGCACTTATGACTGTCTTCTTTATCA
TGCAACTCGTAGGACAGGTGCCGGCAGCGCTCTTCCGCTTCTCGCTCACTGACTCGCTGCG
CTCGGTGCTTCGGCTGCGGCGAGCGGTATCAGCTCACTCAAAGGCGGTAATACGGTTATCCA
CAGAATCAGGGGATAACGCAGGAAAGAACATGTGAGCAAAAGGCCAGCAAAAGGCCAGGAAC
CGTAAAAAGGCCGCGTGTGCTGGCGTTTTTCCATAGGCTCCGCCCCCTGACGAGCATCACAA
AAATCGACGCTCAAGTCAGAGGTGGCGAAACCCGACAGGACTATAAAGATACCAGGCGTTTC
CCCCGGAAGCTCCCTCGTGCGCTCTCCTGTTCCGACCCCTGCCGCTTACCGGATACCTGTCC
GCCTTTCTCCCTTCGGGAAGCGTGGCGCTTTCTCAATGCTCACGCTGTAGGTATCTCAGTTC
GGTGTAGGTGCTTCGCTCCAAGCTGGGCTGTGTGCACGAACCCCCGTTTACGCCCCAGCGCT
GCGCCTTATCCGGTAACCTATCGTCTTGAGTCCAACCCGCTAAGACACGACTTATCGCCACTG
GCAGCAGCCACTGGTAACAGGATTAGCAGAGCGAGGTATGTAGGCGGTGCTACAGAGTTCTT
GAAGTGGTGGCCTAACTACGGCTACACTAGAAGGACAGTATTTGGTATCTGCGCTCTGCTGA
AGCCAGTTACCTTCGGAAAAAGAGTTGGTAGCTCTTGATCCGGCAAAACAAACCCGCTGGT
AGCGGTGGTTTTTTTTGTTTGAAGCAGCAGATTACGCGCAGAAAAAAGGATCTCAAGAAGA
TCCTTTGATCTTTTCTACGGGGTCTGACGCTCAGTGGAACGAAAACTCACGTTAAGGGATTT
TGGTCATGAGATTATCAAAAAGGATCTTACCTAGATCTTTTAAATTAAAAATGAAGTTTT
AAATCAATCTAAAGTATATATGAGTAAACTTGGTCTGACAGTTACCAATGCTTAATCAGTGA
GGCACCTATCTCAGCGATCTGTCTATTTGCTTCATCCATAGTTGCCTGACTCCCCGTCGTGT
AGATAACTACGATACGGGAGGGCTTACCATCTGGCCCCAGTGCTGCAATGATACCGCGAGAC
CCACGCTCACCGGCTCCAGATTTATCAGCAATAAACCAGCCAGCCGGAAGGGCCGAGCGCAG
AAGTGGTCCCTGCAACTTTATCCGCCTCCATCCAGTCTATTAATTGTTGCCGGGAAGCTAGAG
TAAGTAGTTCGCCAGTTAATAGTTTGCAGCAACGTTGTTGCCATTGCTACAGGCATCGTGGTG
TCACGCTCGTCGTTTGGTATGGCTTCATTCAGCTCCGTTCCCAACGATCAAGGCGAGTTAC
ATGATCCCCCATGTTGTGCAAAAAAGCGTTAGCTCCTTCGTTCCCTCCGATCGTTGTCAGAA
GTAAGTTGGCCGAGTGTATCACTCATGGTTATGGCAGCACTGCATAATTCTCTTACTGTC
ATGCCATCCGTAAGATGCTTTTCTGTGACTGGTGAGTACTCAACCAAGTCATTCTGAGAATA
GTGTATGCGGCGACCGAGTTGCTCTTGGCCGGCGTCAATACGGGATAATACCGCGCCACATA
GCAGAACTTTAAAAGTGCTCATCATTTGGAACCGTTCTTCGGGGCGAAAACTCTCAAGGATC
TTACCGCTGTTGAGATCCAGTTCGATGTAACCCACTCGTGACCCCACTGATCTTCAGCATC
TTTTACTTTTACCAGCGTTTTCTGGGTGAGCAAAAAACAGGAAGGCAAAATGCGGCAAAAAAGG
GAATAAGGGCGACACGGAAATGTTGAATACTCATACTCTTCTTTTTTCAATATTATTGAAGC
ATTTATCAGGGTTATTGTCTCATGAGCGGATACATATTTGAATGTATTTAGAAAAATAACA
AATAGGGGTTCCGCGCACATTTCCCCGAAAAGTGCCACCTGACGCGCCCTGTAGCGGCGCAT
TAAGCGCGGCGGGTGTGGTGGTTACGCGCAGCGTGACCGCTACACTTGCCAGCGCCCTAGCG
CCGCTCCTTTGCTTTTCTTCCCTTCTTCTCGCCACGTTTCGCCGGCTTTCCCCGTCAAGC
TCTAAATCGGGGCTCCCTTTAGGGTTCCGATTTAGTGCTTTACGGCACCTCGACCCCAAAA
AAGTTGATTAGGGTGTGTTTACGTTAGTGGGTCATCGCCCTGATAGACGGTTTTTTCGCCCT
TTGACGTTGGAGTCCACGTTCTTTAATAGTGGACTCTTGTTCCAACTGGAACAACACTCAA
CCCTATCTCGGTCTATTCTTTTGTATTTATAAGGGATTTTGCCGATTTCCGGCCTATTGGTTAA
AAAATGAGCTGATTTAAACAAAAATTTAACGCGAATTTTAAACAAAATATTAACGTTTACAATT
TCCCATTCGCCATTTCAGGCTGCGCAACTGTTGGGAAGGGCGATCGGTGCGGGCCTCTTCGCT
ATTACGCCAGCCCAAGCTACCATGATAAGTAAGTAATATTAAGGTACGGGAGGTACTTGGAG
CGGCCGCAATAAAATATCTTTATTTTATTACATCTGTGTGTGGTTTTTTGTGTGAATCGA

TAGTACTAACATACGCTCTCCATCAAAACAAAACGAAACAAAACAAACTAGCAAAATAGGCT
GTCCCCAGTGCAAGTGCAGGTGCCAGAACATTTCTCTATCGATAGGTACCGAGCTCTTACGC
GTGCTAGCCCTCGAGCAGGATCTATACATTGAATCAATATTGGCAATTAGCCATATTAGTCA
TTGGTTATATAGCATAAATCAATATTGGCTATTGGCCATTGCATACGTTGTATCTATATCAT
AATATGTACATTTATATTGGCTCATGTCCAATATGACCGCCATGTTGACATTGATTATTGAC
TAGTTATTAATAGTAATCAATTACGGGGTCATTAGTTCATAGCCCATATATGGAGTTCCGCG
TTACATAACTTACGGTAAATGGCCCGCCTGGCTGACCGCCCAACGACCCCCGCCCATTGACG
TCAATAATGACGTATGTTCCCATAGTAACGCCAATAGGGACTTTCCATTGACGTCAATGGGT
GGAGTATTTACGGTAAACTGCCCACTTGGCAGTACATCAAGTGTATCATATGCCAAGTCCGC
CCCCATTGACGTCAATGACGGTAAATGGCCCGCCTGGCATTATGCCCAGTACATGACCTTA
CGGGACTTTCTACTTGGCAGTACATCTACGTATTAGTCATCGCTATTACCATGGTGATGCG
GTTTTGGCAGTACATCAATGGGCGTGATAGCGGTTTGACTCACGGGGATTTCCAAGTCTCC
ACCCCATGACGTCAATGGGAGTTTGTTTTGGCACCAAAATCAACGGGACTTTCCAAAATGT
CGTAACAACCTCCGCCCCATTGACGCAAAATGGGCGGTAGGCGGTACGGTGGGAGGTCTATAT
AAGCAGAGCTCGTTTAGTGAACCGTCAGATCGCCTGGAGACGCCATCCACGCTGTTTTGACC
TCCATAGAAGACACCGGGACCGATCCAGCCTCCCCTCGAAGCTCGACTCTAGGGGGCTCGAGA
TCCCCGGGTACCGGTGCGCCACCATGGTGAGCAAGGGCGAGGAGCTGTTACCGGGGTGGTGC
CCATCCTGGTTCGAGCTGGACGGCGACGTAAACGGCCACAAGTTCAGCGTGTCCGGCGAGGGC
GAGGGCGATGCCACCTACGGCAAGCTGACCCTGAAGTTCATCTGCACCACCGGCAAGCTGCC
CGTGCCCTGGCCCCACCCTCGTGACCACCCTGACCTACGGCGTGCAGTGCTTCAGCCGCTACC
CCGACCACATGAAGCAGCACGACTTCTTCAAGTCCGCCATGCCCGAAGGCTACGTCCAGGAG
CGCACCATCTTCTTCAAGGACGACGGCAACTACAAGACCCGCGCCGAGGTGAAGTTCGAGGG
CGACACCCTGGTGAACCGCATCGAGCTGAAGGGCATCGACTTCAAGGAGGACGGCAACATCC
TGGGGCACAAGCTGGAGTACAAC TACAACAGCCACAACGTCTATATCATGGCCGACAAGCAG
AAGAACGGCATCAAGGTGAAC TCAAGATCCGCCACAACATCGAGGACGGCAGCGTGCAGCT
CGCCGACCACTACCAGCAGAACACCCCCATCGGCGACGGCCCCGTGCTGCTGCCCCACAACC
ACTACCTGAGCACCCAGTCCGCCCTGAGCAAAGACCCCAACGAGAAGCGCGATCACATGGTC
CTGCTGGAGTTCGTGACCGCCGCGGGATCACTCTCGGCATGGACGAGCTGTACAAGTAAAG
CGGCCGCTCGAGCATGCAT

Fig. 14

p-12.0-lys-LSPIFNMM-CMV-pur-attB (SEQ ID NO: 7)

GGGCTGCAGGAATTCGATTGCCGCCTTCTTTGATATTCACCTCTGTTGTATTTTCATCTCTTCT
TGCCGATGAAAGGATATAACAGTCTGTATAACAGTCTGTGAGGAAATACTTGGTATTTCTTC
TGATCAGTGTTTTTTATAAGTAATGTTGAATATTGGATAAGGCTGTGTGTCCTTTGTCTTGGG
AGACAAAGCCCCACAGCAGGTGGTGGTTGGGTGGCAGCTCAGTGACAGGAGAGGTTTTT
TTGCCCTGTTTTTTTTTTTTTTTTTTTTTAAAGTAAGGTGTTCTTTTTTCTTAGTAAATTTT
CTACTGGACTGTATGTTTTGACAGGTGAGAAACATTTCTTCAAAGAAGAACCCTTTTGAAA
CTGTACAGCCCTTTTCTTTCATTCCCTTTTGTCTTCTGTGCCAATGCCTTTGGTTCTGATT
GCATTATGGAAAACGTTGATCGGAACCTTGAGGTTTTTATTTATAGTGTGGCTTGAAAGCTTG
GATAGCTGTTGTTACACGAGATACCTTATTAAGTTTAGGCCAGCTTGATGCTTTATTTTTTC
CCTTTGAAGTAGTGAGCGTTCTCTGGTTTTTTTCTTTGAAACTGGTGAGGCTTAGATTTTTT
CTAATGGGATTTTTTACCTGATGATCTAGTTGCATACCCAAATGCTTGTAATGTTTTCTTA
GTTAACATGTTGATAACTTCGGATTTACATGTTGTATATACTTGTCTATCTGTGTTTCTAGTA
AAAATATATGGCATTATATAGAAATACGTAATTCCTGATTTCTTTTTTTTTTATCTCTATGCT
CTGTGTGTACAGGTCAAACAGACTTCACTCCTATTTTTATTTATAGAATTTTATATGCAGTC
TGTCGTTGGTTCTTGTGTTGTAAGGATACAGCCTTAAATTTCTAGAGCGATGCTCAGTAAG
GCGGGTTGTACATGGGTTCAAATGTAAACCGGCACGTTTGGCTGCTGCCTTCCCAGATC
CAGGACACTAACTGCTTCTGCACCTGAGGTATAAATCGCTTCAGATCCAGGGAAGTGACAGA
TCCACGTGCATATTCTTAAAGAAGAATGAATACTTTCTAAATATTTTTGGCATAGGAAGCAA
GCTGCATGGATTTGTTTTGGGACTTAAATTATTTTTGGTAACGGAGTGCATAGGTTTTAAACAC
AGTTGCAGCATGCTAACGAGTCACAGCGTTTATGCAGAAGTGATGCCTGGATGCCTGTTGCA
GCTGTTTACGGCACCTGCTTGCAGTGAGCATTTGCAGATAGGGGTGGGTGCTTTGTGTCTGTG
TTCCACACGCTGCCACACAGCCACCTCCCGGAACACATCTCACCTGCTGGGTACTTTTCAA
ACCATTCTAGCAGTAGTAGATGAGTTACTATGAAACAGAGAAGTTCTCAGTTGGATATTCT
CATGGGATGTCCTTTTTTCCCATGTTGGGCAAAGTATGATAAAGCATCTCTATTTGTAAATTA
TGCACTTGTAGTTCTGAATCCTTTCTATAGCACCACTTATTGCAGCAGGTGTAGGCTCTG
GTGTGGCCTGTGTCTGTGCTTCAATCTTTTAAAGCTTCTTTGAAATACACTGACTTGATTG
AAGTCTCTTGAAGATAGTAAACAGTACTTACCTTTGATCCCAATGAAATCGAGCATTTTCAGT
TGTAAGAAATTCGCGCTATTTCATACCATGTAATGTAATTTTACACCCCAAGTGCTGACACT
TTGGAATATATTCAAGTAATAGACTTTGGCCTCACCTCTTGTGTACTGTATTTGTAAATAG
AAAATATTTTTAACTGTGCATATGATTATTACATTATGAAAGAGACATTCTGCTGATCTTCA
AATGTAAGAAAATGAGGAGTGCGTGTGCTTTTATAAATACAAGTGATTGCAAAATTAGTGCAG
GTGTCTTTAAAAAAGTAATATAAAGGACCAGGTGTTTTACAAGTGAAAT
ACATTCCTATTTGGTAAACAGTTACATTTTATGAAGATTACCAGCGCTGCTGACTTTCTAA
ACATAAGGCTGTATTGTCTTCTGTACCATTGCAATTCCTCATTTCCCAATTTGCACAAGGAT
GTCTGGGTAAACTATTCAAGAAATGGCTTTGAAATACAGCATGGGAGCTTGTCTGAGTTGGA
ATGCAGAGTTGCACTGCAAAATGTCAGGAAATGGATGTCTCTCAGAAATGCCCAACTCCAAAG
GATTTTATATGTGTATATAGTAAGCAGTTTCTGATTCCAGCAGGCCAAAGAGTCTGCTGAA
TGTGTGTTGCCGAGACCTGTATTTCTCAACAAGGTAAGATGGTATCCTAGCAACTGCGGA
TTTTAATACATTTTTCAGCAGAAGTACTTAGTTAATCTCTACCTTTAGGGATCGTTTCATCAT
TTTTAGATGTTTAACTTGAATACTGCATACTTTTAGCTTTTCATGGGTTCTTTTTTTTCAG
CCTTTAGGAGACTGTTAAGCAATTTGCTGTCCAACCTTTTGTGTTGGTCTTAAACTGCAATAG
TAGTTTACCTTGTATTGAAGAAATAAAGACCATTTTTATATTAAAAAATACTTTTGTCTGTC
TTCATTTGACTTGTCTGATATCCTTGCAGTGCCCATTTATGTGAGTTCTGTGATATTCAG
ACATCAAACTTAACGTGAGCTCAGTGGAGTTACAGCTGCGGTTTTGATGCTGTTATTTATTT
CTGAAACTAGAAATGATGTTGTCTTCATCTGCTCATCAAACACTTCATGCAGAGTGTAAGGC
TAGTGAGAAATGCATACATTTATTGATACTTTTTTAAAGTCAACTTTTTATCAGATTTTTTT
TTCATTTGAAATATATTGTTTTCTAGACTGCATAGCTTCTGAATCTGAAATGCAGTCTGAT
TGGCATGAAGAAGCACAGCACTCTTCATCTTACTTAACTTCATTTTGGAAATGAAGGAAGTT
AAGCAAGGGCACAGGTCCATGAAATAGAGACAGTGCGCTCAGGAGAAAGTGAACCTGGATTT
CTTTGGCTAGTGTCTAAATCTGTAGTGAGGAAAGTAACACCCGATTCTTGAAGGGCTCC
AGCTTTAATGCTTCCAAATGAAGGTGGCAGGCACTTGGCCACTGGTTATTTACTGCTATTA
TGTCTCAGTTTTCGCAGCTAACCTGGCTTCTCCACTATTGAGCATGGACTATAGCCTGGCTTC
AGAGGCCAGGTGAAGGTTGGGATGGGTGGAAGGAGTGCTGGGCTGTGGCTGGGGGACTGTG

GGGACTCCAAGCTGAGCTTGGGGTGGGCAGCACAGGGAAAAGTGTGGGTAACTATTTTAAAG
TACTGTGTTGCAAACGTCTCATCTGCAAATACGTAGGGTGTGTACTCTCGAAGATTAACAGT
GTGGGTTCAGTAATATATGGATGAATTCACAGTGAAGCATTCAAGGGTAGATCATCTAACG
ACACCAGATCATCAAGCTATGATTGGAAGCGGTATCAGAAGAGCGAGGAAGGTAAGCAGTCT
TCATATGTTTTCCCTCCACGTAAAGCAGTCTGGGAAAGTAGCACCCCTTGAGCAGAGACAAG
GAAATAATTGAGGAGCATGTGCTAGGAGAACTTCTTGCTGAATCTACTTGCAAGAGCTTT
GATGCCTGGCTTCTGGTGCCCTCTGCAGCACCTGCAAGGCCAGAGCCTGTGGTGAGCTGGA
GGGAAAGATTCTGCTCAAGTCCAAGCTTCAGCAGGTCATTGTCTTTGCTTCTTCCCCCAGCA
CTGTGCAGCAGAGTGGAAGTATGTGCAAGCCTCCTGTCCACTACCTGTTGCTGCAGGCAGA
CTGCTCTCAGAAAAAGAGAGCTAACTCTATGCCATAGTCTGAAGGTAAAAATGGGTTTTAAAA
AAGAAAACACAAAGGCAAAACCGGCTGCCCATGAGAAGAAAGCAGTGGTAAACATGGTAGA
AAAGGTGCAGAAGCCCCCAGGCAGTGTGACAGGCCCTCCTGCCACCTAGAGGCGGGAACAA
GCTTCCCTGCCTAGGGCTCTGCCCGCAAGTGCCTGTTTCTTTGGTGGGTTTTGTTTGGCGT
TTGGTTTTGAGATTTAGACACAAGGGAAGCTGAAAGGAGGTGTTGGGCACCTATTTTGGTTT
GTAAAGCCTGTACTTCAAATATATATTTTGTGAGGGAGTGTAGCGAATTGGCCAATTTAAAA
TAAAGTTGCAAGAGATTGAAGGCTGAGTAGTTGAGAGGGTAACACGTTTAATGAGATCTTCT
GAACTACTGCTTCTAAACACTTGTTTGAGTGGTGTGAGACCTTGGATAGGTGAGTGCTCTTGT
TACATGTCTGATGCACCTTGTCTTGTCTTTTCCATCCACATCCATGCATTCCACATCCACGCA
TTTGTCACTTATCCCATATCTGTCTATCTGACATACCTGTCTCTTCGTCACTTGGTCAGAA
GAAACAGATGTGATAATCCCCAGCCGCCCAAGTTTGAGAAGATGGCAGTTGCTTCTTTCCC
TTTTTCCCTGCTAAGTAAGGATTTTTCTCCTGGCTTTGACACCTCACGAAATAGTCTTCTGCC
TTACATTCTGGGCATTATTTCAAATATCTTTGGAGTGCCTGTCTCAAGTTTGTGTCTTCC
TACTCTTAGAGTGAATGCTCTTAGAGTGAAAGAGAAGGAAGAGAAGATGTTGGCCGCAGTTC
TCTGATGAACACACCTCTGAATAATGGCCAAAGGTGGGTGGGTTTTCTCTGAGGAACGGGCAG
CGTTTGCTCTGAAAGCAAGGAGCTCTGCGGAGTTGCAGTTATTTTGCAACTGATGGTGGA
CTGGTGCTTAAAGCAGATTCCCTAGGTTCCCTGCTACTTCTTTTCTTCTTGGCAGTCAGTT
TATTTCTGACAGACAAACAGCCACCCCACTGCAGGCTTAGAAAGTATGTGGCTCTGCCTGG
GTGTTCTGCACTCTGCCCTGGTGAAAGGGGATTAAACGGGCACCATTCATCCCAAACAGG
ATCCTCATTCATGGATCAAGCTGTAAGGAACCTGGGCTCCAACCTCAAACATTAATTGGAG
TACGAATGTAATTAACACTGCATTCTCGCATTCCTAAGTCATTTAGTCTGGACTCTGCAGCA
TGTAGGTGGGCAGCTCCCACTTTCTCAAAGACCACTGATGGAGGAGTAGTAAAAATGGAGAC
CGATTGAGAACAAACCAACGGAGTGTTGCCGAAGAACTGATGGAAATAATGCATGAATTGTG
TGGTGGACATTTTTTTTTAAATACATAAACTACTTCAAATGAGGTGCGAGAAGGTCAAGTGT
TATTAGCAGCCATAAAACAGGTGAGCGAGTACCATTTTTTCTCTACAAGAAAAACGATTCTG
AGCTCTGCGTAAAGTATAAGTTCTCCATAGCGGCTGAAGCTCCCCCTGGCTGCCTGCCATCT
CAGCTGGAGTGCAGTGCCATTTCTTGGGGTTTTCTCTCACAGCAGTAATGGGACAATACTTC
ACAAAAATTCTTTCTTTCTTCTGTCTGTTGGGATCCCTACTGTGCCCTCCTGGTTTTTACGTTA
CCCCCTGACTGTTCCATTGAGCGTTTTGGAAAGAGAAAAAGAAATTTGGAAATAAAACATGTC
TACGTTATCACCTCCTCCAGCATTTTGGTTTTTTAATTATGTCAATAACTGGCTTAGATTTGG
AAATGAGAGGGGGTTGGGTGTATTACCGAGGAACAAAGGAAGGCTTATATAAACTCAAGTCT
TTTATTTAGAGAACTGGCAAGCTGTCAAAAACAAAAAGGCCTTACCACCAAATTAAGTGAAT
AGCCGCTATAGCCAGCAGGGCCAGCACGAGGGATGGTGCACCTGCTGGCACTATGCCACGGCC
TGCTTGTGACTCTGAGAGCAACTGCTTTGGAAATGACAGCACTTGGTGCAATTTCTTTTGT
TCAGAATGCGTAGAGCGTGTGCTTGGCGACAGTTTTTCTAGTTAGGCCACTTCTTTTTCTCT
TCTCTCCTCATCTCCTAAGCATGTCTCCATGCTGGTAATCCAGTCAAGTGAACGTTCAA
CAATGAATCCATCACTGTAGGATTCTCGTGGTGATCAAATCTTTGTGTGAGGTCTATAAAAT
ATGGAAGCTTATTTATTTTTCTGTTCTTCCATATCAGTCTTCTCTATGACAAATTCACATCCAC
CACAGCAAATTAAGGTGAAGGAGGCTGGTGGATGAAGAGGGTCTTCTAGCTTTACGTTCT
TCCTTGCAAGGCCACAGGAAAAATGCTGAGAGCTGTAGAATACAGCCTGGGGTAAGAAGTTCA
GTCTCCTGCTGGGACAGCTAACCGCATCTTATAACCCCTTCTGAGACTCATCTTAGGACCAA
ATAGGGTCTATCTGGGGTTTTTGTTCCTGCTGTTCTCTCTGGAAGGCTATCTCACTATTTCA
CTGCTCCCACGGTTACAAACCAAGATACAGCCTGAATTTTTTCTAGGCCACATTACATAAA
TTTGACCTGTTACCAATATTGTTCTCTATATAGTTATTTCTTCCCCACTGTGTTTAAACCC
TTAAGGCATTGAGAACAACTAGAATCATAGAATGTTTTGGATTGGAAGGGGCCTTAAACATC

ATCCATTTCCAACCCTCTGCCATGGGCTGCTTGCCACCCACTGGCTCAGGCTGCCAGGGCC
CCATCCAGCCTGGCCTTGAGCACCTCCAGGGATGGGGCACCCACAGCTTCTCTGGGCAGCCT
GTGCCAACACCTCACCCTCTCTGGGTAAAGAATTCTCTTTTAACATCTAATCTAAATCTCT
TCTCTTTTAGTTTAAAGCCATTCTCTTTTCCCGTTGCTATCTGTCCAAGAAATGTGTATT
GGTCTCCCTCCTGCTTATAAGCAGGAAGTACTGGAAGGCTGCAGTGAGGTCTCCCCACAGCC
TTCTCTTCTCCAGGCTGAACAAGCCCAGCTCCTTCAGCCTGTCTTCGTAGGAGATCATCTTA
GTGGCCCTCCTCTGGACCCATTCCAACAGTTCCACGGCTTTCTTGTTGGAGCCCCAGGTCTGG
ATGCAGTACTTCAGATGGGGCCTTACAAAGGCAGAGCAGATGGGGACAATCGCTTACCCCTC
CCTGCTGGCTGCCCTGTTTTGATGCAGCCCAGGGTACTGTTGGCCTTTCAGGCTCCCAGAC
CCCTTGCTGATTTGTGTCAAGCTTTTCATCCACCAGAACCCACGCTTCCTGGTTAATACTTC
TGCCCTCACTTCTGTAAAGCTTGTTTCAGGAGACTTCCATTCTTTAGGACAGACTGTGTTACA
CCTACCTGCCCTATTCTTGCAATATATACATTTTCAGTTTCAGTTTCCTGTAAACAGGACAGAAT
ATGTATTCCTCTAACAAAAATACATGCAGAAATTCCTAGTGCCATCTCAGTAGGGTTTTTCATG
GCAGTATTAGCACATAGTCAATTTGCTGCAAGTACCTTCCAAGCTGCGGCCCTCCCATAAATC
CTGTATTTGGGATCAGTTACCTTTTGGGGTAAGCTTTTGTATCTGCAGAGACCCTGGGGGTT
CTGATGTGCTTCAGCTCTGCTCTGTTCTGACTGCACCATTTTCTAGATCACCCAGTTGTTCC
TGTACAACCTTCCTTGCTCCTCCATCCTTTCCAGCTTGATCTTTGACAAATACAGGCCTATT
TTTGTGTTTGCTTCAGCAGCCATTTAATTCTTCAGTGTCTCTTGTCTGTTGATGCCACTG
GAACAGGATTTTCAGCAGTCTTGCAAAGAACATCTAGCTGAAAACCTTCTGCCATTCAATAT
TCTTACCAGTTCTTCTTGTTTGAGGTGAGCCATAAATTACTAGAACTTCGTCACTGACAAGT
TTATGCATTTTATTACTTCTATTATGTACTTACTTTGACATAAACACAGACACGCACATATTT
TGCTGGGATTTCCACAGTGTCTCTGTGTCTTCACATGGTTTTTACTGTCTACTTCCGTTAT
AACCTTGGCAATCTGCCCAGCTGCCCATCACAAGAAAAGAGATTCCTTTTTTTATTACTTCTC
TTCAGCCAATAAACAAAATGTGAGAAAGCCCAAACAAGAACTTGTGGGGCAGGCTGCCATCAA
GGGAGAGACAGCTGAAGGGTTGTGTAGCTCAATAGAATTAAGAAATAATAAAGCTGTGTGAG
ACAGTTTTGCCTGATTTATACAGGCACGCCCCAAGCCAGAGAGGCTGTCTGCCAAGGCCACC
TTGCAGTCTTGGTTTGTAAGATAAGTCATAGGTAACCTTTTCTGGTGAATTGCGTGAGAAAT
CATGATGGCAGTTCTTGCTGTTTACTATGGTAAGATGCTAAAAATAGGAGACAGCAAAGTAAC
ACTTGCTGCTGTAGGTGCTCTGCTATCCAGACAGCGATGGCACTCGCACACCAAGATGAGGG
ATGCTCCCAGCTGACGGATGCTGGGGCAGTAACAGTGGGTCCCATGCTGCCTGCTCATTAGC
ATCACCTCAGCCCTCACCAGCCCATCAGAAGGATCATCCCAAGCTGAGGAAAGTTGCTCATC
TTCTTCACATCATCAAACCTTTGGCCTGACTGATGCCTCCCGGATGCTTAAATGTGGTCACT
GACATCTTTATTTTTCTATGATTTCAAGTCAGAACTCCGGATCAGGAGGGAACACATAGTG
GGAATGTACCCTCAGCTCCAAGGCCAGATCTTCTTCAATGATCATGCTAGTCTTAGGAA
GGTGTGTGTGTGTGAATGTAGAATTGCCTTTGTTATTTTTTCTTCTGCTGTGAGGAACATT
TTGAATACCAGAGAAAAAGAAAAGTGCTCTTCTTGGCATGGGAGGAGTTGTACACTTGCAA
AATAAAGGATGCAGTCCCAAATGTTCAATCTCAGGGTCTGAAGGAGGATCAGAAACTGTG
TATACAATTTTCAGGCTTCTCTGAATGCAGCTTTTGAAAGCTGTTCTTGGCCGAGGCAGTACT
AGTCAGAAACCCTCGGAAACAGGAACAAATGTCTTCAAGGTGCAGCAGGAGGAAACACCTTGC
CCATCATGAAAGTGAATAACCACTGCCGCTGAAGGAATCCAGCTCCTGTTTGAGCAGGTGCT
GCACACTCCACACTGAAACAACAGTTCAATTTTTATAGGACTTCCAGGAAGGATCTTCTTCT
TAAGCTTCTTAATTATGGTACATCTCCAGTTGGCAGATGACTATGACTACTGACAGGAGAAT
GAGGAACTAGCTGGGAATATTTCTGTTTGACCACCATGGAGTCACCCATTTCTTTACTGGTA
TTTGGAATAATAATTTCTGAATTGCAAAGCAGGAGTTAGCGAAGATCTTCATTTCTTCCATG
TTGGTGACAGCACAGTTCTGCTATGAAAGTCTGCTTACAAGGAAGAGGATAAAAAATCATAG
GGATAATAAATCTAAGTTTGAAGACAATGAGGTTTTAGCTGCATTTGACATGAAGAAATTGA
GACCTCTACTGGATAGCTATGGTATTTACGTGCTTTTTGCTTAGTTACTTATTGACCCAG
CTGAGGTCAAGTATGAACTCAGGTCTCTCGGGCTACTGGCATGGATTGATTACATACAACCTG
TAATTTTTAGCAGTGATTTAGGGTTTATGAGTACTTTTGCAGTAAATCATAGGGTTAGTAATG
TTAATCTCAGGGAAAAAAGCCAAACCCTGACAGACATCCAGCTCAGGTGGAAATC
AAGGATCACAGCTCAGTGCAGTCCCAGAGAACACAGGGACTCTTCTCTTAGGACCTTTATGT
ACAGGGCCTCAAGATAACTGATGTTAGTCAGAAAGCTTTCCATTCTGGCCACAGTTCAGCTG
AGGCAATCCTGGAATTTTCTCTCCGCTGCACAGTTCCAGTCATCCAGTTTGTACAGTTCTG
GCATTTTTTGGGTGAGGCCGTGATCCAAGGAGCAGAAGTTCCAGCTATGGTCAGGGAGTGCC

TGACCGTCCCAACTCACTGCACTCAAACAAAGGCGAAACCACAAGAGTGGCTTTTGTGTTGAAA
TTGCAGTGTGGCCAGAGGGGCTGCACCAGTACTGGATTGACCACGAGGCAACATTAATCCT
CAGCAAGTGCAATTTGCAGCCATTAAATTGAACTAACTGATACTACAATGCAATCAGTATCA
ACAAGTGGTTTTGGCTTGGAAGATGGAGTCTAGGGGCTCTACAGGAGTAGCTACTCTCTAATG
GAGTTGCATTTTGAAGCAGGACACTGTGAAAAGCTGGCCTCCTAAAGAGGGCTGCTAAACATT
AGGGTCAATTTTCCAGTGCACCTTCTGAAGTGTCTGCAGTTCCTCATGCAAAGCTGCCAAA
CATAGCACTTCCAATTGAATACAATTATATGCAGGCGTACTGCTTCTTGCCAGCACTGTCTCT
TCTCAAATGAACTCAACAAACAATTTCAAAGTCTAGTAGAAAAGTAACAAGCTTTGAATGTCA
TTAAAAAGTATATCTGCTTTCAGTAGTTTCACTTATTTATGCCCACTAGAAACATCTTGTAC
AAGCTGAACACTGGGGCTCCAGATTAGTGGTAAAACCTACTTTATACAATCATAGAATCATA
GAATGGCCTGGGTTGGAAGGGACCCCAAGGATCATGAAGATCCAACACCCCGCCACAGGCA
GGGCCACCAACCTCCAGATCTGGTACTAGACCAGGCAGCCAGGGCTCCATCCAACCTGGCC
ATGAACACCTCCAGGGATGGAGCATCCACAACCTCTCTGGGCAGCCTGTGCCAGCACCTCAC
CACCTCTCTGTGAAGAACTTTTCCCTGACATCCAATCTAAGCCTTCCCTCCTTGAGGTAG
ATCCACTCCCCCTTGTGCTATCACTGTCTACTCTTGTAAAAAGTTGATTCTCTCCTTTTTG
GAAGGTTGCAATGAGGTCTCCTTGCAAGCTTCTTCTCTTCTGCAGGATGAACAAGCCCAGCT
CCCTCAGCCTGTCTTTATAGGAGAGGTGCTCCAGCCCTCTGATCATCTTTGTGGCCCTCCTC
TGGACCCGCTCCAAGAGCTCCACATCTTTCTGTACTGGGGGCCCCAGGCCTGAATGCAGTA
CTCCAGATGGGGCTCAAAGAGCAGAGTAAAGAGGGACAATCACCTTCCCTCACCTGTCTGG
CCAGCCCTCTTCTGATGGAGCCCTGGATACAACTGGCTTTCTGAGCTGCAACTTCTCCTTAT
CAGTTCCACTATTAAAAACAGGAACAATACACAGGTGCTGATGGCCAGTGCAGAGTTTTCAC
CACTTCTTCATTTCCGGTAGATCTTAGATGAGGAACGTTGAAGTTGTGCTTCTGCGTGTGCTT
CTTCTCTCAAATACTCCTGCCTGATACCTCACCCACCTGCCACTGAATGGCTCCATGGC
CCCCTGCAGCCAGGGCCCTGATGAACCCGGCACTGCTTCAGATGCTGTTTAAATAGCACAGTA
TGACCAAGTTGCACCTATGAATACACAAACAATGTGTTGCATCCTTCAGCACTTGAGAAGAA
GAGCCAAATTTGCATTGTGAGGAAATGGTTTAGTAATTCTGCCAATTAAAACTGTTTATCT
ACCATGGCTGTTTTATGGCTGTTAGTAGTGGTACACTGATGATGAACAATGGCTATGCAGT
AAAATCAAGACTGTAGATATTGCAACAGACTATAAAATTCCTCTGTGGCTTAGCCAATGTGG
TACTTCCCACATTGTATAAGAAATTTGGCAAGTTTAGAGCAATGTTTGAAGTGTGAGGAAAT
TTCTGTATACTCAAGAGGGCGTTTTTGACAACCTGTAGAACAGAGGAATCAAAGGGGGTGGG
AGGAAGTTAAAGAAGAGGCAGGTGCAAGAGAGCTTGCAGTCCCGCTGTGTGTACGACACTG
GCAACATGAGGTCTTTGCTAATCTTGGTGCTTTGCTTCCTGCCCTGGCTGCCTTAGGGTGC
GATCTGCCTCAGACCCACAGCCTGGGCAGCAGGAGGACCCTGATGCTGCTGGCTCAGATGAG
GAGAATCAGCCTGTTTAGCTGCCTGAAGGATAGGCACGATTTTGGCTTTCCTCAAGAGGAGT
TTGGCAACCACTTTGAGAGGCTGAGACCATCCCTGTGCTGCACGAGATGATCAGCAGATC
TTTAACCTGTTTAGCACCAAGGATAGCAGCGCTGCTTGGGATGAGACCTGCTGGATAAGTT
TTACACCGAGCTGTACCAGCAGCTGAACGATCTGGAGGCTTGCCTGATCCAGGGCGTGGGCG
TGACCGAGACCCCTCTGATGAAGGAGGATAGCATCCTGGCTGTGAGGAAGTACTTTAGAGG
ATCACCTGTACCTGAAGGAGAAGAAGTACAGCCCTGCGCTTGGGAAGTCTGTAGGGCTGA
GATCATGAGGAGCTTTAGCCTGAGCACCAACCTGCAAGAGAGCTTGAGGTCTAAGGAGTAAA
AAGTCTAGAGTCGGGGCGGCCGCCGCTTCGAGCAGACATGATAAGATACATTGATGAGTTT
GGACAAACCACAACCTAGAATGCAGTGAATAAATGCTTTATTTGTGAAATTTGTGATGCTAT
TGCTTTATTTGTAACCATTTATAAGCTGCAATAAACAAGTTAACAACAACAATTGCATTCATT
TTATGTTTCAGGTTTCAAGGGGAGGTGTGGGAGGTTTTTTAAAGCAAGTAAAACCTCTACAAA
TGTGGTAAATCGATAAGGATCCGTCGACCGATGCCCTTGAGAGCCTTCAACCCAGTCAGCT
CCTTCCGGTGGGCGCGGGGCATGACTATCGTCGCCGCACTTATGACTGTCTTCTTTATCATG
CAACTCGTAGGACAGGTGCCGGCAGCGCTCTTCCGCTTCCCTCGCTCACTGACTCGCTGCGCT
CGGTCTGTCGGCTGCGGCGAGCGGTATCAGCTCACTCAAAGGCGGTAAACGGTTATCCACA
GAATCAGGGGATAACGACAGGAAGAACATGTGAGCAAAAGGCCAGCAAAAGGCCAGGAACCG
TAAAAAGGCCGCGTTGTGCTGGCGTTTTTCCATAGGCTCCGCCCCCTGACGAGCATCACAAAA
ATCGACGCTCAAGTCAGAGGTGGCGAAACCCGACAGGACTATAAAGATACCAGGCGTTTCCC
CCTGGAAGCTCCCTCGTGCCTCTCCTGTTCCGACCCCTGCCGCTTACCGGATACCTGTCCGC
CTTTCTCCTTCCGGAAGCGTGGCGTTTTCTCAATGCTCACGCTGTAGGTATCTCAGTTCCG
TGTAGGTCGTTCCGCTCCAAGCTGGGCTGTGTGCACGAACCCCGTTTCCAGCCCGACCGCTGC

GCCTTATCCGGTAACTATCGTCTTGAGTCCAACCCGGTAAGACACGACTTATCGCCACTGGC
AGCAGCCACTGGTAACAGGATTAGCAGAGCGAGGTATGTAGGCGGTGCTACAGAGTTCTTGA
AGTGGTGGCCTAACTACGGCTACACTAGAAGGACAGTATTTGGTATCTGCGCTCTGCTGAAG
CCAGTTACCTTTCGGAAAAAGAGTTGGTAGCTCTTGATCCGGCAAACAAACCACCGCTGGTAG
CGGTGGTTTTTTTTGTTTGCAAGCAGCAGATTACGCGCAGAAAAAAGGATCTCAAGAAGATC
CTTTGATCTTTTTCTACGGGGTCTGACGCTCAGTGGAAACGAAAACCTCACGTTAAGGGATTTTG
GTCATGAGATTATCAAAAAGGATCTTCACCTAGATCCTTTTTAAATTAAAAATGAAGTTTTAA
ATCAATCTAAAGTATATATGAGTAACTTGGTCTGACAGTTACCAATGCTTAATCAGTGAGG
CACCTATCTCAGCGATCTGTCTATTTTCGTTTCATCCATAGTTGCCTGACTCCCCGTCGTGTAG
ATAACTACGATACGGGAGGGCTTACCATCTGGCCCCAGTGCTGCAATGATACCGCGAGACCC
ACGCTCACCGGCTCCAGATTTATCAGCAATAAACCAGCCAGCCGGAAGGGCCGAGCGCAGAA
GTGGTCTGCAACTTTATCCGCCTCCATCCAGTCTATTAATTGTTGCCGGGAAGCTAGAGTA
AGTAGTTCCGCAGTTAATAGTTTGCGCAACGTTGTTGCCATTGCTACAGGCATCGTGGTGTCT
ACGCTCGTCTGTTTGGTATGGCTTCATTCAGCTCCGGTCCCAACGATCAAGGCGAGTTACAT
GATCCCCCATGTTGTGCAAAAAAGCGGTTAGCTCCTTCGGTCTCCGATCGTTGTCAGAAGT
AAGTTGGCCGCGAGTGTTATCACTCATGGTTATGGCAGCACTGCATAATTCTCTTACTGTCAT
GCCATCCGTAAGATGCTTTTCTGTGACTGGTGAGTACTCAACCAAGTCATTCTGAGAATAGT
GTATGCGGCGACCGAGTTGCTCTTGCCCGGCGTCAATACGGGATAATACCGCGCCACATAGC
AGAACCTTAAAAGTGCTCATCATTGGAAAACGTTCTTCGGGGCGAAAACCTCTCAAGGATCTT
ACCGCTGTTGAGATCCAGTTCGATGTAAACCCACTCGTGCACCCAAGTATCTTCAGCATCTT
TTACTTTTACCAGCGTTTCTGGGTGAGCAAAAACAGGAAGGCAAAATGCCGCAAAAAGGGA
ATAAGGGCGACACGGAAATGTTGAATACTCATACTCTTCCTTTTTCAATATTATTGAAGCAT
TTATCAGGGTTATTGTCTCATGAGCGGATACATATTTGAATGTATTTAGAAAAATAAACAAA
TAGGGGTTCCGCGCACATTTCCCCGAAAAGTGCCACCTGACGCGCCCTGTAGCGGCGCATTA
AGCGCGGCGGGTGTGGTGGTTACGCGCAGCGTGACCGCTACACTTGCCAGCGCCCTAGCGCC
CGCTCCTTTTCGCTTTCTTCCCTTCTTCTCGCCACGTTTCGCCGGCTTTCCCCGTCAGCTC
TAAATCGGGGGCTCCCTTTAGGGTTCCGATTTAGTGCTTTACGGCACCTCGACCCCAAAAA
CTTGATTAGGGTGATGGTTACGTTAGTGGGCCATCGCCCTGATAGACGGTTTTTTTCGCCCTTT
GACGTTGGAGTCCACGTTCTTTAATAGTGGACTCTTGTTCCAACTGGAACAACACTCAACC
CTATCTCGGTCTATTCTTTTGATTTATAAGGGATTTTGCCGATTTTCGCCCTATTGGTTAAAA
AATGAGCTGATTTAACAAAAATTTAACGCGAATTTTAAACAAAATATTAACGTTTACAATTTTC
CCATTCGCCATTACAGGCTGCGCAACTGTTGGGAAGGGCGATCGGTGCGGGCCTCTTCGCTAT
TACGCCAGCCCAAGCTACCATGATAAGTAAGTAATATTAAGGTACGGGAGGTACTTGGAGCG
GCCGCTCTAGACTAGTGGATCCCCCGGCGCAATAAAATATCTTTATTTTCATTACATCTGT
TGTGTTGGTTTTTTTTGTGTGAATCGATAGTACTAACATACGCTCTCCATCAAAACAAAACGAA
ACAAAACAACTAGCAAAATAGGCTGTCCCCAGTGCAAGTGACAGGTGCCAGAACATTTCTCT
ATCGATAGGTACCGAGCTCTTACGCGTGCTAGCCCTCGAGCAGGATCTATACATTGAATCAA
TATTGGCAATTAGCCATATTAGTTCATTGGTTATATAGCATAAATCAATATTGGCTATTGGCC
ATTGCATACGTTGTATCTATATCATAATATGTACATTTATATTGGCTCATGTCCAATATGAC
CGCCATGTTGACATTGATTATTGACTAGTTATTAATAGTAATCAATTACGGGGTTCATTAGTT
CATAGCCCATATATGGAGTTCCGCGTTACATAACTTACGGTAAATGGCCCGCCTGGCTGACC
GCCCAACGACCCCCGCCATTGACGTCAATAATGACGTATGTTCCCATAGTAACGCCAATAG
GGACTTTCCATTGACGTCAATGGGTGGAGTATTTACGGTAAACTGCCCACTTGGCAGTACAT
CAAGTGTATCATATGCCAAGTCCGCCCTTATTGACGTCAATGACGGTAAATGGCCCGCCTG
GCATTATGCCAGTACATGACCTTACGGGACTTTCCTACTTGGCAGTACATCTACGTATTAG
TCATCGCTATTACCATGGTGATGCGGTTTTTGGCAGTACATCAATGGGCGTGGATAGCGGTTT
GACTCACGGGGATTTCCAAGTCTCCACCCCATGACGTCAATGGGAGTTTGTTTTGGCACCA
AAATCAACGGGACTTTCCAAAATGTGTAACAACCTCCGCCCATTGACGCAAAATGGGCGGTA
GGCGTGTACGGTGGGAGGTCTATATAAGCAGAGCTCGTTTGTAGTGAACCGTCAGATCGCCTGG
AGACGCCATCCACGCTGTTTTGACCTCCATAGAAGACACCGGGACCGATCCAGCCTCCCCTC
GAAGCTCGACTCTAGGGGCTCGAGATCTGCGATCTAAGTAAGCTTGATGCCTGCAGGTCCG
CCGCCACGACCGGTGCCGCCACCATCCCCTGACCCACGCCCCCTGACCCCTCACAAGGAGACG
ACCTTCCATGACCGAGTACAAGCCACGGTGCGCCTCGCCACCCGCGACGACGTCCCCCGGG
CCGTACGCACCCCTCGCCGCCGCGTTTCGCCGACTACCCCGCCACGCGCCACACCGTCGACCCG

GACCGCCACATCGAGCGGGTCACCGAGCTGCAAGAACTCTTCCTCACGCGCGTCGGGGCTCGA
CATCGGCAAGGTGTGGGTGCGCGACGACGGCGCCGCGGTGGCGGTCTGGACCACGCCGGAGA
GCGTCGAAGCGGGGGCGGTGTTCCGCCGAGATCGGGCCGCGCATGGCCGAGTTGAGCGGTTCC
CGGCTGGCCGCGCAGCAACAGATGGAAGGCCTCCTGGCGCCGCACCGGCCCAAGGAGCCCGC
GTGGTTCTTGCCACCGTCGGCGTCTCGCCCCGACCACCAGGGCAAGGGTCTGGGCAGCGCCG
TCGTGCTCCCCGGAGTGGAGGCGGCCGAGCGCGCCGGGGTGCCCGCCTTCCTGGAGACCTCC
GCGCCCCGCAACCTCCCCTTCTACGAGCGGCTCGGCTTCACCGTCACCGCCGACGTCGAGGT
GCCCCAAGGACCGCGCACCTGGTGATGACCCGCAAGCCCGGTGCCTGACGCCCCGCCACG
ACCCGCAGCGCCCGACCGAAAGGAGCGCACGACCCCATGGCTCCGACCGAAGCCGACCCGGG
CGGCCCCCGCCGACCCCGCACCCGCCCCCGAGGCCACCGACTCTAGAGTCGGGGCGGCCGGC
CGCTTCGAGCAGACATGATAAGATACATTGATGAGTTTGGACAAACCACAACCTAGAATGCAG
TGAAAAAATGCTTTATTTGTGAAATTTGTGATGCTATTGCTTTATTTGTAACCATTTATAAG
CTGCAATAAACAAGTTAACAACAACAATTGCATTCATTTTATGTTTCAGGTTTCAAGGGGAGG
TGTGGGAGGTTTTTTAAAGCAAGTAAAACCTCTACAAATGTGGTAAAATCGATAAGGATCAA
TTCGGCTTCAGGTACCGTCGACGATGTAGGTCACGGTCTCGAAGCCGCGGTGCGGGTGCCAG
GGCGTGCCCTTGGGCTCCCCGGGCGCGTACTCCACCTCACCCATCTGGTCCATCATGATGAA
CGGGTCGAGGTGGCGGTAGTTGATCCCGGCGAACGCGCGGCGCACCGGGAAGCCCTCGCCCT
CGAAACCGCTGGGCGCGGTGGTCACGGTGAGCACGGGACGTGCGACGGCGTCGGCGGGTGCG
GATACGCGGGGCAGCGTCAGCGGGTCTCGACGGTCACGGCGGGCATGTGACAGCCGAATT
GATCCGTCGACCGATGCCCTTGAGAGCCTTCAACCCAGTCAGCTCCTTCCGGTGGGCGCGGG
GCATGACTATCGTCGCCGCACTTATGACTGTCTTCTTTATCATGCAACTCGTAGGACAGGTG
CCGGCAGCGCTCTTCCGCTTCCTCGCTCACTGACTCGCTGCGCTCGGTGCTTCGGCTGCGGC
GAGCGGTATCAGCTCACTCAAAGGCGGTAATACGGTTATCCACAGAATCAGGGGATAACGCA
GGAAAGAACATG

Fig. 15

pOM IFN-Ins-CMV-pur-attB (SEQ ID NO: 8)

GGCCGCCACCGCGGTGGAGCTCCAATTCGCCCTATAGTGAGTCGTATTACAATTCACTGGCC
GTCGTTTTTACAACGTCGTGACTGGGAAAACCCCTGGCGTTACCCAACCTAATCGCCTTGCAGC
ACATCCCCCTTTTCGCCAGCTGGCGTAATAGCGAAGAGGCCCGCACCGATCGCCCTTCCCAAC
AGTTGCGCAGCCTGAATGGCGAATGGGACGCGCCCTGTAGCGGCGCATTAAGCGCGGCGGGT
GTGGTGGTTACGCGCAGCGTGACCGCTACACTTGCAGCGCCCTAGCGCCCGCTCCTTTTCGC
TTTCTTCCCTTCTTTCTCGCCACGTTTCGCCGGCTTTCCCGTCAAGCTCTAAATCGGGGGC
TCCCTTTAGGGTTCGGATTTAGTGCTTTACGGCACCTCGACCCCAAAAACTTGATTAGGGT
GATGGTTCACGTAGTGGGCCATCGCCCTGATAGACGGTTTTTTCGCCCTTTGACGTTGGAGTC
CACGTTCTTTAATAGTGGACTCTTGTTCCAAACTGGAACAACACTCAACCCTATCTCGGTCT
ATTCTTTTGATTTATAAGGGATTTTGCCGATTTTCGGCCTATTGGTTAAAAAATGAGCTGATT
TAACAAAAATTTAACGCGAATTTTAACAAAATATTAACGCTTACAATTTAGGTGGCACTTTT
CGGGGAAATGTGCGCGGAACCCCTATTGTTTATTTTCTAAATACATTCAAATATGTATCC
GCTCATGAGACAATAACCCCTGATAAATGCTTCAATAATATTGAAAAAGGAAGAGTATGAGTA
TTCAACATTTCCGTGTGCGCCCTTATTCCCTTTTTTTCGGCATTTTGCCTTCTGTTTTTGCT
CACCCAGAAACGCTGGTGAAAGTAAAGATGCTGAAGATCAGTTGGGTGCACGAGTGGGTTA
CATCGAAGTGGATCTCAACAGCGGTAAGATCCTTGAGAGTTTTTCGCCCGGAAGAACGTTTTTC
CAATGATGAGCACTTTTAAAGTTCTGCTATGTGGCGCGGTATTATCCCGTATTGACGCCGGG
CAAGAGCAACTCGGTGCGCGCATACACTATTCTCAGAATGACTTGGTTGAGTACTCACCAGT
CACAGAAAAGCATCTTACGGATGGCATGACAGTAAGAGAATTATGCAAGTGTGCCATAACCA
TGAGTGATAACATGCGGCCAACTTACTTCTGACAACGATCGGAGGACCGAAGGAGCTAACC
GCTTTTTTGCACAACATGGGGGATCATGTAACCTCGCCTTGATCGTTGGGAACCGGAGCTGAA
TGAAGCCATACCAAACGACGAGCGTGACACCACGATGCCTGTAGCAATGGCAACAACGTTGC
GCAAACTATTAACTGGCGAACTACTTACTCTAGCTTCCCGGCAACAATTAATAGACTGGATG
GAGGCGGATAAAGTTGCAGGACCACTTCTGCGCTCGGCCCTTCCGGCTGGCTGGTTTTATTGC
TGATAAATCTGGAGCCGGTGAGCGTGGGTCTCGCGGTATCATTTGCAGCACTGGGGCCAGATG
GTAAGCCCTCCCGTATCGTAGTTATCTACACGACGGGGAGTCAGGCAACTATGGATGAACGA
AATAGACAGATCGCTGAGATAGGTGCCTCACTGATTAAGCATTTGGTAACTGTCAGACCAAGT
TTACTCATATATACTTTAGATTGATTTAAACTTCATTTTTAATTTAAAGGATCTAGGTGA
AGATCCTTTTTTGATAATCTCATGACCAAAATCCCTTAACGTGAGTTTTTCGTTCCACTGAGCG
TCAGACCCCGTAGAAAAGATCAAAGGATCTTCTTGAGATCCTTTTTTTCTGCGCGTAATCTG
CTGCTTGCAAACAAAAAACCACCGCTACCAGCGGTGGTTTTGTTTGCCGGATCAAGAGCTAC
CAACTCTTTTTCCGAAGGTAACCTGGCTTCAGCAGAGCGAGATACCAATACTGTCCTCTA
GTGTAGCCGTAGTTAGGCCACCACTTCAAGAACTCTGTAGCACCGCCTACATACCTCGCTCT
GCTAATCCTGTTACCAGTGGCTGCTGCCAGTGGCGATAAGTCGTGTCTTACCGGGTTGGACT
CAAGACGATAGTTACCGGATAAGGCGCAGCGGTGCGGCTGAACGGGGGGTTCTGTGCACACAG
CCCAGCTTGGAGCGAACGACCTACACCGAACTGAGATACCTACAGCGTGAGCTATGAGAAAG
CGCCACGCTTCCCGAAGGGAGAAAGGCGGACAGGTATCCGGTAAGCGGCAGGGTCGGAACAG
GAGAGCGCACGAGGGAGCTTCCAGGGGGAAACGCCTGGTATCTTTATAGTCTGTCTGGGTTT
CGCCACCTCTGACTTGAGCGTCGATTTTTGTGTGCTCGTCAGGGGGGCGGAGCCTATGGAA
AAACGCCAGCAACGCGGCCCTTTTTACGGTTCTTGGCCTTTTGCTGGCCTTTTGCTCACATGT
TCTTTCTGCGTTATCCCCTGATTCTGTGGATAACCGTATTACCGCCTTTGAGTGAGCTGAT
ACCGCTCGCCGACGCCGAACGACCGAGCGCAGCGAGTCAGTGAGCGAGGAAGCGGAAGAGCG
CCCAATACGCAAACCGCCTCTCCCCGCGCGTTGGCCGATTCATTAATGCAGCTGGCACGACA
GGTTTCCCGACTGGAAAGCGGGCAGTGAGCGCAACGCAATTAATGTGAGTTAGCTCACTCAT
TAGGCACCCCAGGCTTTACACTTTATGCTTCCGGCTCGTATGTTGTGTGGAATTTGTGAGCGG
ATAACAATTTACACAGGAAACAGCTATGACCAATGATTACGCCAAGCTCGAAATTAACCCCTC
ACTAAAGGGGAACAAAAGCTGGGTACCGGGCCCCCCCCCTCGACTAGAGGGACAGCCCCCCCCCA
AAGCCCCCAGGGATGTAATTACGTCCCTCCCCCGCTAGGGGGCAGCAGCGAGCCGCCCGGGG
CTCCGCTCCGGTCCGGCGCTCCCCCGCATCCCCGAGCCGGCAGCGTGCGGGGACAGCCCGG
GCACGGGGAAGGTGGCACGGGATCGCTTCTCTGAACGCTTCTCGCTGCTCTTTGAGCCTG
CAGACACCTGGGGGGATACGGGGAAAAAGCTTTAGGCTGAAAGAGAGATTTAGAATGACAGA
ATCATAGAACGGCCTGGGTGCAAAGGAGCACAGTGCTCATCCAGATCCAACCCCTGCTAT
GTGCAGGGTCATCAACCAGCAGCCCAGGCTGCCAGAGCCACATCCAGCCTGGCCTTGAATG

CCTGCAGGGATGGGGCATCCACAGCCTCCTTGGGCAACCTGTTTCAGTGCGTCACCACCCCTCT
GGGGGAAAAACTGCCTCCTCATATCCAACCCAAACCTCCCCTGTCTCAGTGTAAGCCATTC
CCCCTTGTCTATCAAGGGGGAGTTTGCTGTGACATTGTTGGTCTGGGGTGACACATGTTTG
CCAATT CAGTG CATCACGAGAGGCAGATCTTGGGGATAAGGAAGTGCAGGACAGCATGGAC
GTGGGACATGCAGGTGTTGAGGGCTCTGGGACACTCTCCAAGTCACAGCGTT CAGAACAGCC
TTAAGGATAAGAAGATAGGATAGAAGGACAAAGAGCAAGTTAAAACCCAGCATGGAGAGGAG
CACAAAAAGGCCACAGACACTGCTGGTCCCTGTGTCTGAGCCTGCATGTTTGATGGTGCTG
GATGCAAGCAGAAGGGGTGGAAGAGCTTGCTGGAGAGATAACAGCTGGGT CAGTAGGACTGG
GACAGGCAGCTGGAGAATTGCCATGTAGATGTT CATAACAATCGTCAAATCATGAAGGCTGGA
AAAGCCCTCCAAGATCCCCAAGACCAACCCCAACCCACCCACCGTGCCCACTGGCCATGTCC
CTCAGTGCCACATCCCCACAGTTCTTCATCACCTCCAGGGACGGTGACCCCCCACCTCCGT
GGGCAGCTGTGCCACTGCAGCACCGCTCTTTGGAGAAGGTAAATCTTGCTAAATCCAGCCCG
ACCCTCCCCTGGCACAACGTAAGGCCATTATCTCTCATCCAACCTCCAGGACGGAGTCAGTGA
GGATGGGGCTCTAGTCGAGGTCGACGGTATCGATAAGCTTGATTAGGCAGAGCAATAGGACT
CTCAACCTCGTGAGTATGGCAGCATGTTAACTCTGCAGCTGGAGTCCAGCGTGGGAAACAATC
TGCCTTGACATGAGTCTTCGTGGGCCAATATTCCCCAACGGTTTTTCTTCAGCTTGTCTTG
TCTCCTAAGCTCTCAAAACACCTTTTTTGGTGAATAAACTCACTTGGCAACGTTTATCTGTCT
TACCTTAGTGTCAGGTTTCATCCCTATTCCCCTTTCTCCTCCTCCGTGTGGTACACAGTGGT
GCACACTGGTTCTTCTGTGTGATGTTCTGCTCTGACAGCCAATGTGGGTAAAGTTCTTCCTGC
CACGTGTCTGTGTTGTTTTCACTTCAAAAAGGGCCCTGGGCTCCCCTTGGAGCTCTCAGGCA
TTTTCTTAATCATCACAGTCACGCTGGCAGGATTAGTCCCTCCTAAACCTTAGAATGACCTG
AACGTGTGCTCCCTCTTTGTAGTCAGTGACGGGAGACGTTTGCCTCAAGATCAGGGTCCATC
TCACCCACAGGGCCATTCCCAAGATGAGGTGGATGTTTTACTCTCAAAAAAGTTTTCTTAT
GTTTGGCTAGAAAGGAGAACTCACTGCCTACCTGTGAATTTCCCTAGTCTGTGCTGC
CACTGCTGCCTGTGCAGCCTGTCCATGGAGGGGGCAGCAACTGCTGTCAAAAGGTGATCC
CACCTGTCTCCACTGAAATGACCTCAGTGCCACGTGTTGTATAGGGTATAAAGTACGGGAG
GGGGATGCCCGGCTCCCTTCAGGGTTGCAGAGCAGAAGTGTCTGTGTATAGAGTGTGCTTA
ATCTATTAATGTAACAGAACAACTTCAGTCTTAGTGTTTTGTGGGCTGGAATTGCCCATGTG
GTAGGGACAGGCCTGTCTAAATCACTGCAATCGCCTATGTTCTGAAGGTATTTGGGAAAGAAA
GGGATTTGGGGGATTGCCTGTGATTGGCTTTAATTGAATGGCAAATCACAGGAAAGCAGTTC
TGCTCAACAGTTGGTTGTTTCAGCCAATTCTTGCAGCCAAAGAGCCGGGTGCCAGCGATAT
AATAGTTGTCACTTGTGTCTGTATGGATGACAGGGAGGTAGGGTGACCTGAGGACCACCTC
CAGCTTCTGCTAGCGTAGGTACAGTCACCACCTCCAGCTCCACACGAGTCCCATCGTGGTTT
ACCAAAGAAACACAATTATTTGGACCAGTTTGGAAAGTCAACCGCTGAATTGTGAGGCTAGA
TTAATAGAGCTGAAGAGCAAATGTTCCCACTTGGAGATACTAGTTGGTATAGTATCAGAG
GAACAGGGCCATAGCACCTCCATGCTATTAGATTCCGGCTGGCATGTACTTTTCAAGATGAT
TTGTAACTAACAAATGGCTTATTGTGCTTGTCTTAAGTCTGTGTCTAATGTAAATGTTCTT
TGGTTTTATATAACCTTCTTGCCATTTGCTCTTCAGGTGTTCTTGCAGAACACTGGCTGCTTT
AATCTAGTTTTAACTGTTGCTTGATTATTCTTAGGGATAAGATCTGAATAAACTTTTTGTGGC
TTTGGCAGACTTTAGCTTGGGCTTAGCTCCACATTAGCTTTTGCTGCCTTTTCTGTGAAGC
TATCAAGATCCTACTCAATGACATTAGCTGGGTGCAGGTGTACCAAATCCTGCTCTGTGGAA
CACATTGTCTGATGATACCGAAGGCAAACGTGAACTCAAAGAGGCACAGAGTTAAGAAGAAG
TCTGTGCAATT CAGAGGAAAAGCCAAAGTGGCCATTAGACACACTTTCATGCAGCATTTGC
CAGTAGGTTTT CATATAAACTACAAAATGGAATAAACCACTACAAATGGGAAAAGCCTGATA
CTAGAATTTAAATATT CACCCAGGCTCAAGGGGTGTTTCATGGAGTAATATCACTCTATAAA
AGTAGGGCAGCCAATTATTCACAGACAAAGCTTTTTTTTTTCTGTGCTGCAGTGCTGTTTTT
CGGCTGATCCAGGGTTACTTATTGTGGGTCTGAGAGCTGAATGATTTCTCCTTGTGT CATGT
TGGTGAAGGAGATATGGCCAGGGGGAGATGAGCATGTTCAAGAGGAAACGTTGCATTTTGGT
GGCTTGGGAGAAAGGTAGAACGATATCAGGTCCATAGTGTCACTAAGAGATCTGAAGGATGG
TTTTTACAGAACAGTTGACTTGGCTGGGTGCAGGCTTGGCTGTAAATGGATGGAAGGATGGAC
AGATGGGTGGACAGAGATTTCTGTGCAGGAGATCATCTCCTGAGCTCGGTGCTTGACAGACT
GCAGATCCATCCATAACCTTCTCCAGCATGAGAGCGCGGGGAGCTTTGGTACTGTT CAGTC
TGCTGCTTGTGCTTCTGGGTGCACAGTGGTGATTTTCTTACTCACACAGGGCAAAAACCT
GAGCAGCTTCAAAGTGAACAGGTGCTCTCATAGGCCATT CAGTTGTCAAGATGAGGTTTTT

[illegible]

GGGCAGCAGAGTGGCTCTGGGTCACTGTGGCACCAATGGGAGGCACCAGTGGGGGTGTGTTT
TGTGCAGGGGGGAAGCATTACAGAATGGGGCTGATCCTGAAGCTTGACAGTCCAAGGCTTTG
TCTGTGTACCCAGTGAATCCTTCTCTGTACATAAAGCCCAGATAGGACTCAGAAATGTA
GTCATTCCAGCCCCCTCTTCTCAGATCTGGAGCAGCACTTGTTTGCAGCCAGTCTCCCC
AAAATGCACAGACCTCGCCGAGTGGAGGGAGATGTAAACAGCGAAGGTTAATTACCTCCTTG
TCAAAAACACTTTGTGGTCCATAGATGTTTCTGTCAATCTTACAAAACAGAACCGAGAGGCA
GCGAGCACTGAAGAGCGTGTTCCTCATGTGAGTTAATGAGACTTGGCAGCTCGCTGTGCAGA
GATGATCCCTGTGCTTCATGGGAGGCTGTAACCTGTCTCCCCATCGCCTTCACACCGCAGTG
CTGTCTGGACACCTCACCTCCATAAGCTGTAGGATGCAGCTGCCCAGGGATCAAGAGACT
TTTCTAAGGCTCTTAGGACTCATCTTTGCCGCTCAGTAGCGTGCAGCAATTACTCATCCCA
ACTATACTGAATGGGTTTCTGCCAGCTCTGCTTGTGTGTCATAAGCATTCTTTCATTTTGC
CTCTAAGTTTCTCTCAGCAGCACCGCTCTGGGTGACCTGAGTGGCCACCTGGAACCCGAGGG
GCACAGCCACCACCTCCCTGTTGCTGCTGCTCCAGGGAATCATGTGCTGCTGGATGGGGGGA
AGCATGAATGTTCTCACCCAGACACCTGGTGTGCAATGGCTGCAGCGTGCCTTCTTGGTAT
GCAGATTGTTTCCAGCCATTACTTGTAGAAATGTGCTGTGGAAGCCCTTTGTATCTCTTTCT
GTGGCCCTTCAGCAAAAGCTGTGGGAAAGCTCTGAGGCTGCTTTCTTGGGTCTGGAGGAAT
TGTATGTTCTCTTTTAAACAAAATTATCTTAGGAGAGAGCACTGTGCAAGCATTGTGCAC
ATAAAACAATTACAGGTTGAAAGGGCTCTCTGGAGGTTTCCAGCCTGACTACTGCTCGAAGCA
AGGCCAGGTTCAAAGATGGCTCAGGATGCTGTGTGCCTTCTGATTATCTGTGCCACCAATG
GAGGAGATTACAGCCACTCTGCTTCCCGTGCCACTCATGGAGAGGAATATTCCTTATATT
CAGATAGAATGTTATCCTTTAGCTCAGCCTTCCCTATAACCCCATGAGGGAGCTGCAGATCC
CCATACTCTCCCCCTTCTCTGGGGTGAAGGCCGTGTCCCCCAGCCCCCTTCCCACCTGTGC
CCTAAGCAGCCCGCTGGCCTCTGCTGGATGTGTGCCTATATGTCAATGCCTGTCTTGCAGT
CCAGCCTGGGACATTTAATTATCACCAGGGTAATGTGGAACCTGTGTCTCTCCCTGCAG
GGTACAAAGTTCTGCACGGGGTCTTTTCCGTTTCAGGAAAACCTTCACTGGTGCTACCTGAAT
CAAGCTCTATTTAATAAGTTTATAAGCACATGGATGTGTTTTCTTAGAGATACGTTTTAATG
GTATCAGTGATTTTTATTTGCTTTGTGCTTACTTCAAACAGTGCCCTTTGGGCAGGAGGTGA
GGGACGGGTCTGCCGTTGGCTCTGCAGTGATTTCTCCAGGCGTGTGGCTCAGGTCAGATAGT
GGTCACTCTGTGGCCAGAAGAAGGACAAAGATGGAAATTGCAGATTGAGTCACGTTAAGCAG
GCATCTTGGAGTGATTTGAGGCAGTTTCATGAAAGAGCTACGACCACTTATTGTTGTTTTCC
CCTTTTACAACAGAAGTTTTTATCAAATAACGTGGCAAAGCCCAGGAATGTTTGGGAAAAG
TGTAAGTTAAATGTTTTGTAAATTCATTTGTGCGAGTGCTACCAGCTAAGAAAAAAGTCTACC
TTTGGTATGGTAGTCTTGCAGAGAATACAACATCAATATTAGTTTGGAAAAAAACACCACCA
CCACCAGAACTGTAATGGAAAAATGTAACCAAGAAATTCTTGGGTAAAGAGAGAAAAGGATG
TCGTATAGTGGCCAGTCCCTGCCAGCTCAGCTGCTCCTCATGCAGTGCCAAAACCTCGTGTTGA
GAAACGTGGCACTGTAAGACGTGTCCCTGCTTTGCTTGGCCACAGATCTCTGCCCTTGTG
CTGACTCTGCACACAAGAGCATTTCCCTGTAGCCAAACAGCGATTAGCCATAAGCTGCACC
TGACTTTGAGGATTAAGAGTTTGAATTAAGTGGATTGCAGCAGGAGATCAGTGGCAGGGTT
GCAGATGAAATCCTTTTCTAGGGGTAGCTAAGGGCTGAGCAACCTGTCCTACAGCACAAAGCC
AAACCAGCCAAGGGTTTTCTCTGTGCTGTTACAGAGGCAGGGCCAGCTGGAGCTGGAGGAGG
TTGTGCTGGGACCTTCTCCCTGTGCTGAGAATGGAGTGATTTCTGGGTGCTGTTCTCTGTGG
CTTGCACTGAGCAGCTCAAGGGAGATCGGTGCTCCTCATGCAGTGCCAAAACCTCGTGTTGA
TGCAGAAAGATGGATGTGCACCTCCCTCCTGCTAATGCAGCCGTGAGCTTATGAAGGCAATG
AGCCCTCAGTGCAGCAGGAGCTGTAGTGCACTCCTGTAGGTGCTAGGGAAAATCTCTGGTTC
CCAGGGATGCATTCATAAGGGCAATATATCTTGAGGCTGCGCCAAATCTTTCTGAAATATTC
ATGCGTGTTCCCTTAATTTATAGAAACAAACACAGCAGAAATAATTATCCAATGCCTCCCCCT
CGAAGGAAACCCATATTTCCATGTAGAAATGTAACCTATATACACAGCCATGCTGCATCC
TTCAGAACGTGCCAGTGCTCATCTCCCATGGCAAAATACTACAGGTATTTCTCACTATGTTGG
ACCTGTGAAAGGAACCATGGTAAGAAACTTCGGTTAAAGGTATGGCTGCAAAACTACTCATA
CCAAAACAGCAGAGCTCCAGACCTCCTCTTAGGAAAGAGCCACTTGGAGAGGGATGGTGTGA
AGGCTGGAGGTGAGAGACAGAGCCTGTCCAGTTTTCTGTCTCTATTTTCTGAAACGTTTTG
CAGGAGGAAAGGACAACCTGTACTTTTCAAGCATAGCTGGTGGCCTCACGTAAATAAGTTCCCC
GAACCTCTGTGTCAATTTGTTCTTAAGATGCTTTGGCAGAACACTTTGAGTCAATTTCGCTTAA
CTGTGACTAGGTCTGTAAATAAGTGCTCCCTGCTGATAAGGTTCAAGTGACATTTTTTAGTGG

TATTTGACAGCATTTTACCTTGCTTTTCAAGTCTTCTACCAAGCTCTTCTATACTTAAGCAGTG
AAACCGCCAAGAAACCCCTTCTTTTATCAAGCTAGTGCTAAATACCATTAACTTCATAGGTT
AGATACGGTGCTGCCAGCTTCACCTGGCAGTGGTTGGTCAGTTCTGCTGGTGACAAAGCCTC
CCTGGCCTGTGCTTTTACCTAGAGGTGAATATCCAAGAATGCAGAACTGCATGGAAAGCAGA
GCTGCAGGCACGATGGTGCTGAGCCTTAGCTGCTTCTGCTGGGAGATGTGGATGCAGAGAC
GAATGAAGGACCTGTCCCTTACTCCCCCTCAGCATTCTGTGCTATTTAGGGTTCTACCAGAGT
CCTTAAGAGGTTTTTTTTTTTTTTTGGTCCAAAAGTCTGTGTTGTTTGGTTTTTGACCACTGAGA
GCATGTGACACTTGTCTCAAGCTATTAACCAAGTGTCAGCCAAAATCAATTGCCTGGGAGA
CGCAGACCATTACCTGGAGGTGAGGACCTCAATAAATATTACCAGCCTCATTGTGCCGCTGA
CAGATTGAGCTGGCTGCTCCGTGTTCCAGTCCAACAGTTTCGGACGCCACGTTTGTATATATT
TGCAGGCAGCCTCGGGGGGACCATCTCAGGAGCAGAGCACCGGCAGCCGCTGCAGAGCCGG
GCAGTACCTCACCATGGCTTTGACCTTTGCCTTACTGGTGGCTCTCCTGGTGCTGAGCTGCA
AGAGCAGCTGCTCTGTGGGCTGCGATCTGCCTCAGACCCACAGCCTGGGCAGCAGGAGGACC
CTGATGCTGCTGGCTCAGATGAGGAGAATCAGCCTGTTTAGCTGCCTGAAGGATAGGCAGCA
TTTTGGCTTTCTCAAGAGGAGTTTGGCAACCAGTTTCAGAAGGCTGAGACCATCCCTGTGC
TGCACGAGATGATCCAGCAGATCTTTAACCTGTTTAGCACCAAGGATAGCAGCGCTGCTTGG
GATGAGACCCTGCTGGATAAGTTTTTACACCGAGCTGTACCAGCAGCTGAACGATCTGGAGGC
TTGCGTGATCCAGGGCGTGGGCGTGACCGAGACCCCTCTGATGAAGGAGGATAGCATCCTGG
CTGTGAGGAAGTACTTTTCAAGGATCACCCCTGTACCTGAAGGAGAAGAAGTACAGCCCCCTGC
GCTTGGGAAGTCGTGAGGGCTGAGATCATGAGGAGCTTTAGCCTGAGCACCAACCTGCAAGA
GAGCTTGAGGTCTAAGGAGTAAAAAGTCTAGAGTCGGGGCGGCCGCGCTTCGAGCAGACA
TGATAAGATACATTGATGAGTTTGGACAAACCACAACCTAGAATGCAGTGAAAAAATGCTTT
ATTTGTGAAATTTGTGATGCTATTGCTTTATTTGTAACCATTATAAGCTGCAATAAACAAGT
TAACAACAACAATTGCATTTCATTTTATGTTTTCAGGTTTCAGGGGGAGGTGTGGGAGGTTTTTT
AAAGCAAGTAAAACCTCTACAAATGTGGTAAAATCGATACCGTTCGACCTCGACTAGAGCGGC
CACTAACATACGCTCTCCATCAAAAACAAAACGAAAACAAAACAACTAGCAAAATAGGCTGTC
CCCAGTCAAGTGACAGGTGCCAGAACATTTCTCTATCGATAGGTACCGAGCTCTTACGCGTG
CTAGCCCTCGAGCAGGATCTATACATTGAATCAATATTGGCAATTAGCCATATTAGTCATTG
GTTATATAGCATAAATCAATATTGGCTATTGGCCATTGCATACGTTGTATCTATATCATAAT
ATGTACATTTATATTGGCTCATGTCCAATATGACCGCCATGTTGACATTGATTATTGACTAG
TTATTAATAGTAATCAATTACGGGGTCAATTAGTTTCATAGCCCATATATGGAGTTCCGCGTTA
CATAACTTACGTTAAATGGCCCCGCTGGCTGACCGCCCAACGACCCCCGCCCATTGACGTCA
ATAATGACGTATGTTCCCATAGTAACGCCAATAGGGACTTTCCATTGACGTCAATGGGTGGA
GTATTTACGGTAAACTGCCCACTTGGCAGTACATCAAGTGTATCATATGCCAAGTCCGCCCC
CTATTGACGTCAATGACGTAATGGCCCCGCTGGCATTATGCCCAGTACATGACCTTACGG
GACTTTCTACTTGGCAGTACATCTACGTATTAGTCATCGCTATTACCATTGGTGATGCGGTT
TTGGCAGTACATCAATGGGCGTGGATAGCGGTTTGAATCACGGGGATTTCCAAGTCTCCACC
CCATTGACGTCAATGGGAGTTTGTGTTTGGCACCAAAATCAACGGGACTTTCCAAAATGTCGT
AACAACTCCGCCCCATTGACGCAATGGGCGGTAGGCGGTGTACGGTGGGAGGTCTATATAAG
CAGAGCTCGTTTAGTGAACCGTTCAGATCGCCTGGAGACGCCATCCACGCTGTTTTGACCTCC
ATAGAAGACACCGGGACCGATCCAGCCTCCCCTCGAAGCTCGACTCTAGGGGCTCGAGATCT
GCGATCTAAGTAAGCTTGCATGCTGAGGTGCGGCCGCCACGACCGGTGCCGCCACCATCCC
CTGACCCACGCCCCCTGACCCCTCACAAGGAGACGACCTTCCATGACCGAGTACAAGCCCACG
GTGCGCCTCGCCACCCGCGACGACGTCCCCCGGGCCGTACGCACCCCTCGCCGCCGCGTTCCG
CGACTACCCCGCCACGCGCCACACCGTTCGACCCGGACCGCCACATCGAGCGGGTCAACGAGC
TGCAAGAACTCTTCTCAGCGCGGTGCGGCTCGACATCGGCAAGGTGTGGGTGCGCGGACGAC
GGCGCCGCGGTGGCGGTCTGGACCACGCCGAGAGCGTGAAGCGGGGGCGGTGTTCCGCGA
GATCGGCCCCGCGCATGGCCGAGTTGAGCGTTTCCCGGCTGGCCGCGCAGCAACAGATGGAAG
GCCTCTGGCGCCGACCGGCCAAGGAGCCCGCGTGGTTCTTGGCCACCGTTCGCGCTCTCG
CCCCGACCACAGGGCAAGGGTCTGGGCAGCGCCGTCGTGCTCCCCGGAGTGGAGGCGGCCGA
GCGCGCCGGGGTGCCCGCTTCTTGGAGACCTCCGCGCCCCGCAACCTCCCCTTCTACGAGC
GGCTCGGCTTCACCGTCAACGCCGACGTGAGGTGCCCGAAGGACCGCGCACCTGGTGATG
ACCCGCAAGCCCGGTGCTGACGCCCCGCCACGACCCGCGAGCGCCGACCGAAAGGAGCGC
ACGACCCCATGGCTCCGACCGAAGCCGACCCGGGCGGCCCGCGACCCCGCACCCGCCCCC

GAGGCCACCGACTCTAGAGTCGGGGCGGCCGGCCGCTTCGAGCAGACATGATAAGATAACAT
TGATGAGTTTGGACAAACCACAACCTAGAATGCAGTGAAAAAATGCTTTATTTGTGAAATTT
GTGATGCTATTGCTTTATTTGTAACCATTATAAGCTGCAATAAACAAGTTAACAACAACAAT
TGCAATTCATTTTATGTTTCAGGTTTCAGGGGGAGGTGTGGGAGGTTTTTTTAAAGCAAGTAAAA
CCTCTACAAATGTGGTAAAATCGATAAGGATCAATTCGGCTTCAGGTACCGTCGACGATGTA
GGTCACGGTCTCGAAGCCGCGGTGCGGGTGCCAGGGCGTGCCCTTGGGCTCCCCGGGCGCGT
ACTCCACCTCACCCATCTGGTCCATCATGATGAACGGGTGAGGTGGCGGTAGTTGATCCCG
GCGAACGCGCGGCGCACCGGGAAGCCCTCGCCCTCGAAACCGCTGGGCGCGGTGGTCACGGT
GAGCACGGGACGTGCGACGGCGTCGGCGGGTGCGGATACGCGGGGCAGCGTCAGCGGGTTCT
CGACGGTCACGGCGGGCATGTGACAGCCGAATTGATCCGTCGACCGATGCCCTTGAGAGCC
TTCAACCCAGTCAGCTCCTTCGGGTGGGCGCGGGGCATGACTATCGTCGCCGCACTTATGAC
TGTCTTCTTTATCATGCAACTCGTAGGACAGGTGCCGGCAGC

Fig. 16

pRSV-C31int (SEQ ID NO: 9)

CTGCATTAATGAATCGGCCAACGCGCGGGGAGAGGCGGTTTGCCTATTGGGCGCTCTTCC
GCTTCCTCGCTCACTGACTCGCTGCGCTCGGTCTGGTCTGGCTGCGGCGAGCGGTATCAGCT
CACTCAAAGGCGGTAATACGGTTATCCACAGAATCAGGGGATAACGCAGGAAAGAACATG
TGAGCAAAAGGCCAGCAAAAGGCCAGGAACCGTAAAAAGGCCGCGTTGCTGGCGTTTTTTC
CATAGGCTCCGCCCCCTGACGAGCATCACAAAATCGACGCTCAAGTCAGAGGTGGCGA
AACCCGACAGGACTATAAAGATACCAGGCGTTCCCCCTGGAAGCTCCCTCGTGCGCTCT
CCTGTTCCGACCCCTGCCGCTTACCGGATACCTGTCCGCTTTCTCCCTTCGGGAAGCGTG
GCGCTTTCTCAATGCTCAGCTGTAGGTATCTCAGTTCGGTGTAGGTCTGTTCTCCAAG
CTGGGCTGTGTGCACGAACCCCCCGTTACGCCCCGACCGCTGCGCCTTATCCGGTAACATAT
CGTCTTGAGTCCAACCCGGTAAGACACGACTTATCGCCACTGGCAGCAGCCACTGGTAAC
AGGATTAGCAGAGCGAGGTATGTAGGCGGTGCTACAGAGTCTTGAAGTGGTGGCCTAAC
TACGGCTACACTAGAAGGACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTTC
GGAAAAAGAGTTGGTAGCTCTTGATCCGGCAAAACAAACCACCGCTGGTAGCGGTGGTTTT
TTTGTGTTGCAAGCAGCAGATTACGCGCAGAAAAAAGGATCTCAAGAAGATCCTTTGATC
TTTTCTACGGGTCTGACGCTCAGTGGAACGAAACTCACGTTAAGGGATTTTGGTCATG
AGATTATCAAAAAGGATCTTCACCTAGATCCTTTTAAATTAAAAATGAAGTTTTAAATCA
ATCTAAAGTATATATGAGTAACTTGGTCTGACAGTTACCAATGCTTAATCAGTGAGGCA
CCTATCTCAGCGATCTGTCTATTTCTGTTTCATCCATAGTTGCCTGACTCCCCGTCGTGTAG
ATAACTACGATACGGGAGGGCTTACCATCTGGCCCCAGTGCTGCAATGATACCGCGAGAC
CCACGCTCACCGGCTCCAGATTTATCAGCAATAAACAGCCAGCCGGAAGGCCGAGCGC
AGAAGTGGTCTTGCAACTTTATCCGCTCCATCCAGTCTATTAAATTGTTGCCGGAAGCT
AGAGTAAGTATTGCGCCAGTTAATAGTTTGGCGAACGTTGTTGCCATTGCTACAGGCATC
GTGGTGTACGCTCGTCTGTTTGGTATGGCTTCATTAGCTCCGGTTCCCAACGATCAAGG
CGAGTTACATGATCCCCCATGTTGTGCAAAAAAGCGGTTAGCTCCTTCGGTCTCCGATC
GTTGTGCAAGTAAGTTGGCCGAGTGTTATCACTCATGTTATGGCAGCACTGCATAAT
TCTCTTACTGTCATGCCATCCGTAAGATGCTTTTCTGTGACTGGTGAGTACTCAACCAAG
TCATTTCTGAGAATAGTGTATGCGGCGACCGAGTTGCTCTTGCCCGCGCTCAATACGGGAT
AATACCGCGCCACATAGCAGAACTTTAAAGTGCTCATCATTTGGAAAACGTTCTTCGGGG
CGAAAACCTCTCAAGGATCTTACCGCTGTTGAGATCCAGTTCGATGTAACCCACTCGTGCA
CCCAACTGATCTTCAGCATCTTTTACTTTTACCAGCGTTTCTGGGTGAGCAAAAACAGGA
AGGCAAAATGCCGCAAAAAAGGGAATAAGGGCGACACGGAAATGTTGAATACTCATACTC
TTCTTTTTTCAATATTATTGAAGCATTTATCAGGGTTATTGTCTCATGAGCGGATACATA
TTTGAATGTATTTAGAAAAATAAACAAATAGGGGTTCCGCGCACATTTCCCCGAAAAGTG
CCACCTGACGTGCGACGGATCGGGAGATCTCCCGATCCCTATGGTCGACTCTCAGTACAA
TCTGCTCTGATGCCGCATAGTTAAGCCAGTATCTGCTCCCTGCTTGTGTGTTGGAGGTG
CTGAGTAGTGCGCGAGCAAAATTTAAGCTACAACAAGGCAAGGCTTGACCGACAATTGCA
TGAAGAATCTGCTTAGGGTTAGGCGTTTTTGGCGCTGCTTCGCGATGTACGGGCCAGATATA
CGCGTGCTAGGGGTCTAGGATCGATTCTAGGAATTCTCTAGCCGCGGTCTAGGGATCCCG
GCGCGTATGGTGCACTCTCAGTACAATCTGCTCTGATGCCGCATAGTTAAGCCAGTATCT
GCTCCCTGCTTGTGTGTTGGAGGTGCTGAGTAGTGCGCGAGCAAAATTTAAGCTACAAC
AAGGCAAGGCTTGACCGACAATTGCATGAAGAATCTGCTTAGGGTTAGGCGTTTTGCGCT
GCTTCGCGATGTACGGGCCAGATATACGCGTATCTGAGGGGACTAGGGTGTGTTTAGGCG
AAAAGCGGGGCTTCGGTTGTACGCGGTTAGGAGTCCCTCAGGATATAGTAGTTTCGCTT
TTGCATAGGGAGGGGGAAATGTAGTCTTATGCAATACACTTGTAGTCTTGCAACATGGTA
ACGATGAGTTAGCAACATGCCTTACAAGGAGAGAAAAAGCACCGTGATGCCGATTGGTG
GAAGTAAGGTGGTACGATCGTGCTTATTAGGAAGGCAACAGACAGGTCTGACATGGATT
GGACGAACCACTGAATTCGCGATTGCAGAGATAATTGTATTTAAGTGCCCTAGCTCGATAC
AATAAACGCCATTTGACCATTCACCACATTGGTGTGCACCTCCAAGCTTGATGCCTGCA
GGTACCGGTCCGGAATTCCCGGGTCGACGAGCTCACTAGTCTGAGGGTCGCCGACATGAC
ACAAGGGGTTGTGACCGGGGTGGACACGTACGCGGGTGCTTACGACCGTCAGTCGCGCGA
GCGCGAGAATTCGAGCGCAGCAAGCCCAGCGACACAGCGTAGCGCCAACGAAGACAAGGC
GGCCGACCTTCAGCGCGAAGTCGAGCGCGACGGGGGCCGTTTCAAGTTCGTGCGGCATTT
CAGCGAAGCGCCGGGCACGTCGGCGTTCCGGACGGCGGAGCGCCCGAGTTTCAACGCAT

CCTGAACGAATGCCGCGCCGGGCGGCTCAACATGATCATTGTCTATGACGTGTGCGGCTT
CTCGCGCCTGAAGGTTCATGGACGCGATTCCGATTGTCTCGGAATTGCTCGCCCTGGGCGT
GACGATTGTTTTCCACTCAGGAAGGCGTCTTCCGCGCAGGGAAACGTCATGGACCTGATTCA
CCTGATTATGCGGCTCGACGCGTTCGCACAAAGAATCTTCGCTGAAGTCGGCGAAGATTCT
CGACACGAAGAACCTTCAGCGCGAATTGGGCGGGTACGTGCGCGGGAAGGCGCCTTACGG
CTTCGAGCTTGTTTCGGAGACGAAGGAGATCACGCGCAACGGCCGAATGGTCAATGTCTGT
CATCAACAAGCTTGCGCACTCGACCACTCCCTTACCGGACCCTTCGAGTTCGAGCCCGA
CGTAATCCGGTGGTGGTGGCGTGAGATCAAGACGCACAAACACCTTCCCTTCAAGCCGGG
CAGTCAAGCCGCCATTACCCGGGCAGCATCACGGGGCTTTGTAAGCGCATGGACGCTGA
CGCCGTGCCGACCCGGGGCGAGACGATTGGGAAGAAGACCGCTTCAAGCGCCTGGGACCC
GGCAACC GTTATGCGAATCCTTCGGGACCCGCGTATTGCGGGCTTCGCCGCTGAGGTGAT
CTACAAGAAGAAGCCGGACGGCACGCCGACCACGAAGATTGAGGGTTACCGCATTACGCG
CGACCCGATCACGCTCCGGCCGGTTCGAGCTTGATTGCGGACCGATCATCGAGCCCGCTGA
GTGGTATGAGCTTCAGGCGTGGTTGGACGGCAGGGGGCGCGGCAAGGGGCTTTCCCGGGG
GCAAGCCATTCTGTCCGCCATGGACAAGCTGTACTGCGAGTGTGGCGCCGTATGACTTC
GAAGCGCGGGGAAGAATCGATCAAGGACTCTTACCGCTGCCGTGCCGGAAGGTGGTTCGA
CCCGTCCGCACCTGGGCAGCACGAAGGCACGTGCAACGTCAGCATGGCGGCACTCGACAA
GTTCTGTTGCGGAACGCATCTTCAACAAGATCAGGCACGCCGAAGGCGACGAAGAGACGTT
GGCGCTTCTGTGGGAAGCCGCCCGACGCTTCGGCAAGCTCACTGAGGCGCCTGAGAAGAG
CGGCGAACGGGCGAACCTTGTTGCGGAGCGCGCCGACGCCCTGAACGCCCTTGAAGAGCT
GTACGAAGACCGCGCGGCAGGCGCGTACGACGGACCCGTTGGCAGGAAGCACTTCCGGAA
GCAACAGGCAGCGCTGACGCTCCGGCAGCAAGGGGCGGAAGAGCGGCTTGCCGAACCTTGA
AGCCGCCGAAGCCCCGAAGCTTCCCTTGACCAATGGTTCCCCGAAGACGCCGACGCTGA
CCCGACCGGCCCTAAGTCGTGGTGGGGGCGCGCTCAGTAGACGACAAGCGCGTGTTCGT
CGGGCTCTTCGTAGACAAGATCGTTGTCACGAAGTCGACTACGGGCAGGGGGCAGGGAAAC
GCCCATCGAGAAGCGCGCTTCGATCACGTGGGCGAAGCCGCCGACCGACGACGACGAAGA
CGACGCCCAGGACGGCACGGAAGACGTAGCGGCGTAGCGAGACACCCGGATCCCTCGAGG
GGCCCTATTCTATAGTGTACCTAAATGCTAGAGCTCGCTGATCAGCCTCGACTGTGCCT
TCTAGTTGCCAGCCATCTGTTGTTTGCCCTCCCCCGTGCCTTCCTTGACCCTGGAAGGT
GCCACTCCCACCTGTCTTTCCTAATAAAATGAGGAAATTGCATCGCATTGTCTGAGTAGG
TGTCATTCTATTCTGGGGGGTGGGGTGGGGCAGGACAGCAAGGGGGAGGATTGGGAAGAC
AATAGCAGGCATGCTGGGGATGCGGTGGGCTCTATGGCTTCTGAGGCGGAAAGAACCAGG
TGCCAGTCATAGCCGAATAGCCTCTCCACCCAAGCGGCCGAGAACCCTGCGTGCAATCC
ACTGGGGGCGCG

Fig. 17

pCR-XL-TOPO-CMV-PUR-attB (SEQ ID NO: 10)

AGCGCCCAATACGCAAACCGCCTCTCCCCGCGCGTTGGCCGATTCAATTAATGCAGCTGGC
ACGACAGGTTTCCCGACTGGAAAGCGGGCAGTGAGCGCAACGCAATTAATGTGAGTTAGC
TCACTCATTAGGCACCCAGGCTTTACACTTTATGCTTCCGGCTCGTATGTTGTGTGGAA
TTGTGAGCGGATAACAATTTACACAGGAAACAGCTATGACCATGATTACGCCAAGCTAT
TTAGGTGACGCGTTAGAATACTCAAGCTATGCATCAAGCTTGGTACCGAGCTCGGATCCA
CTAGTAACGCGCCGCGAGTGTGCTGGAATTGCGCCTTGGCCGCAATAAAATATCTTTATTT
TCATTACATCTGTGTGTTGGTTTTTTGTGTGAATCGATAGTACTAACATACGCTCTCCAT
CAAAACAAAACGAAACAAAACAACTAGCAAAATAGGCTGTCCCAGTGCAAGTGCAGGT
GCCAGAACATTTCTCTATCGATAGGTACCGAGCTCTTACGCGTGCTAGCCCTCGAGCAGG
ATCTATACATTGAATCAATATTGGCAATTAGCCATATTAGTCATTGGTTATATAGCATAA
ATCAATATTGGCTATTGGCCATTGCATACGTTGTATCTATATCATAATATGTACATTTAT
ATTGGCTCATGTCCAATATGACCGCCATGTTGACATTGATTATTGACTAGTTATTAATAG
TAATCAATTACGGGGTCATTAGTTTCATAGCCCATATATGGAGTTCCGCGTTACATAAECTT
ACGGTAAATGGCCCGCCTGGCTGACCGCCCAACGACCCCGCCCATTTGACGTCAATAATG
ACGTATGTTCCCATAGTAACGCCAATAGGGACTTTCATTGACGTCAATGGGTGGAGTAT
TTACGGTAAACTGCCCACTTGGCAGTACATCAAGTGTATCATATGCCAAGTCCGCCCCCT
ATTGACGTCAATGACGGTAAATGGCCCGCCTGGCATTATGCCAGTACATGACCTTACGG
GACTTTCTTACTTGGCAGTACATCTACGTATTAGTCATCGCTATTACCATGTGATCGG
TTTTGACGTACATCAATGGGCGTGGATAGCGGTTTGACTACGGGGATTTCCAAGTCTC
CACCCCATTTGACGTCAATGGGAGTTTGTTTTGGCACCAAATCAACGGGACTTTCCAAAA
TGTGCTAACAACTCCGCCCCATTGACGCAAATGGGCGGTAGGCGGTGTACGGTGGGAGGTC
TATATAAGCAGAGCTCGTTTGTGAACCGTCAGATCGCCTGGAGACGCCATCCACGCTGT
TTTGACCTTCATAGAAGACACCGGGACCGATCCAGCCTCCCTCGAAGCTCGACTCTAGG
GGCTCGGATCTGCGATCTAAGTAAGCTTGCATGCCCTGCAGGTGGCCGCCACGACCGGT
GCCGCCACCATCCCCTGACCCACGCCCCTGACCCCTCACAAGGAGACGACCTTCCATGAC
CGAGTACAAGCCACGGTGCGCCCTCGCCACCCGCGACGACGTCCCCCGGGCCGTACGCAC
CCTCGCCGCCGCGTTGCGCGACTACCCCGCCACGCGCCACACCGTCGACCCGGACCGCCA
CATCGAGCGGGTCACCGAGCTGCAAGAACTCTTCTCACGCGCGTCGGGCTCGACATCGG
CAAGGTGTGGGTGCGGGACGACGGCGCGCGGTGGCGGTCTGGACCACGCCGGAGAGCGT
CGAAGCGGGGGCGGTGTTGCGCCGAGATCGGCCCCGCGCATGGCCGAGTTGAGCGGTTCCCG
GCTGGCCGCGCAGCAACAGATGGAAGGCCTCCTGGCGCCGACCCGGCCCAAGGAGCCCGC
GTGGTTCTTGGCCACCGTCGGCGTCTCGCCCGACCACCAGGGCAAGGGTCTGGGCGAGCGC
CGTCGTGCTCCCCGAGTGGAGGCGCCGAGCGCGCCGGGGTGCCCGCTTCTGGAGAC
CTCCGCGCCCCGCAACCTCCCTTCTACGAGCGGCTCGGCTTACCGTCAACGCGGACGT
CGAGGTGCCCCGAAGGACCGCGCACCTGGTGCATGACCCGCAAGCCCGGTGCCTGACGCCC
GCCCCACGACCCGCGAGCGCCCGACCGAAAGGAGCGCACGACCCCATGGCTCCGACCGAAG
CCGACCCGGGCGGCCCCCGGACCCCGCACCCGCCCCCGAGGCCCACCGACTCTAGAGTC
GGGGCGGCCGCGCTTCGAGCAGACATGATAAGATACATTGATGAGTTTGGACAAACCA
CAACTAGAATGCAGTGAAAAAATGCTTTATTTGTGAAATTTGTGATGCTATTGCTTTAT
TTGTAACCATTATAAGCTGCAATAAACAAGTTAACAACAACAATTGCATTCATTTTATGT
TTCAGGTTTACGGGGGAGGTGTGGGAGGTTTTTTAAAGCAAGTAAACCTCTACAAATGTG
GTAAATCGATAAGGATCAATTCGGCTTCAGGTACCGTCGACGATGTAGGTACGGTCTC
GAAGCCGCGGTGCGGGTGCCAGGGCGTGCCCTTGGGCTCCCGGGCGCGTACTCCACCTC
ACCCATCTGGTCCATCATGATGAACGGGTGAGGTGGCGGTAGTTGATCCCGGCGAACGC
GCGGCGCACCGGGAAGCCCTCGCCCTCGAAACCGCTGGGCGCGGTGGTCACGGTGAGCAC
GGGACGTGCGACGGCGTCCGGCGGTGCGGATACGCGGGGCGAGCGTCAGCGGGTTCTCGAC
GGTCACGGCGGGCATGTGACAGCCGAATTGATCCGTGACCGATGCCCTTGAGAGCCTT
CAACCCAGTCAGCTCCTTCCGGTGGGCGCGGGGCGATGACTATCGTCGCCGCACTTATGAC
TGTCTTCTTTATCATGCAACTCGTAGGACAGGTGCCGGCAGCGCTCTTCCGCTTCTCGC
TCACTGACTCGCTGCGCTCGGTGCTTGGCTGCGGCGAGCGGTATCAGCTCACTCAAAGG
CGGTAATACGGTTATCCACAGAATCAGGGGATAACGCAGGAAAGAACATGAAGGGCGAAT
TCTGCAGATATCCATCACACTGGCGGCGCGCTCGAGCATGCATCTAGAGGGCCAAATTCGC
CCTATGTGAGTCGTATTACAATTCAGTGGCCGTCGTTTTTACAACGTCGTGATGGGAAA
ACCTTGGCGTTACCCAACTTAATCGCCTTGCAGCACATCCCCCTTTCGCCAGCTGGCGTA
ATAGCGAAGAGGCGCCGACCGATCGCCCTTCCCAACAGTTGCGCAGCCTATACGTACGGC
AGTTTAAGGTTTACACCTATAAAAGAGAGAGCCGTTATCGTCTGTTTGTGGATGTACAGA
GTGATATTATTGACACGCCGGGGCGACGGATGGTGATCCCCCTGGCCAGTGCACGTCTGC

TGTCAGATAAAGTCTCCCGTGAACCTTTACCCGGTGGTGTCATATCGGGGATGAAAGCTGGC
GCATGATGACCACCGATATGGCCAGTGTGCCGGTCTCCGTTATCGGGGAAGAAGTGGCTG
ATCTCAGCCACCGCGAAAATGACATCAAAAACGCCATTAACCTGATGTTCTGGGGAATAT
AAATGTCAGGCATGAGATTATCAAAAAGGATCTTCACCTAGATCCTTTTCACGTAGAAAG
CCAGTCCGCAGAAACGGTGCTGACCCCGGATGAATGTCAGCTACTGGGCTATCTGGACAA
GGGAAAACGCAAGCGCAAAGAGAAAAGCAGGTAGCTTGCAGTGGGCTTACATGGCGATAGC
TAGACTGGGCGGTTTTATGGACAGCAAGCGAACCGGAATTGCCAGCTGGGGCGCCCTCTG
GTAAGGTTGGGAAGCCCTGCAAAGTAAACTGGATGGCTTTCTCGCCGCCAAGGATCTGAT
GGCGCAGGGGATCAAGCTCTGATCAAGAGACAGGATGAGGATCGTTTCGCATGATTGAAC
AAGATGGATTGCACGCAGGTTCTCCGGCCGCTTGGGTGGAGAGGCTATTCTGGCTATGACT
GGGCACAACAGACAATCGGCTGCTCTGATGCCGCCGTGTTCCGGCTGTGACGCGAGGGGC
GCCCGGTTCTTTTTGTCAAGACCGACCTGTCCGGTGCCCTGAATGAACTGCAAGACGAGG
CAGCGCGGCTATCGTGGCTGGCCACGACGGGCGTTCTTGCAGCTGTGCTCGACGTTG
TCACTGAAGCGGAAGGGACTGGCTGCTATTGGGCGAAGTGCCGGGGCAGGATCTCCTGT
CATCTCACCTTGCTCCTGCCGAGAAAGTATCCATCATGGCTGATGCAATGCCGCGGCTGC
ATACGCTTGATCCGGCTACCTGCCCATTCGACCACCAAGCGAAACATCGCATCGAGCGAG
CACGTACTCGGATGGAAGCCGGTCTTGTGTCGATCAGGATGATCTGGACGAAGAGCATCAGG
GGCTCGCGCCAGCCGAACTGTTCCGCCAGGCTCAAGGCGAGCATGCCCGACGGCGAGGATC
TCGTCTGACCCATGGCGATGCCTGCTTGCCGAATATCATGGTGGAAAATGGCCGCTTTT
CTGGATTTCATCGACTGTGGCCGGCTGGGTGTGGCGGACCGCTATCAGGACATAGCGTTGG
CTACCCGTGATATTGCTGAAGAGCTTGGCGGCGAATGGGCTGACCGCTTCTCGTGCTTT
ACGGTATCGCCGCTCCCGATTGCGAGCGCATCGCCTTCTATCGCCTTCTTGACGAGTTCT
TCTGAATTATTAACGCTTACAATTTCTGATGCGGTATTTTCTCCTTACGCATCTGTGCG
GTATTTTACACCGCATACAGGTGGCACTTTTCGGGGAAATGTGCGCGGAACCCCTATTTG
TTTTATTTTTCTAAATACATTCAAATATGTATCCGCTCATGAGACAATAACCTTGATAAAT
GCTTCAATAATAGCACGTGAGGAGGGCCACCATTGGCCAAAGTTGACCAGTGCCGTTCCGGT
GCTCACCGCGCGGACGTGCGCCGGAGCGGTGAGTTCTGGACCGACCGGCTCGGGTTCTC
CCGGGACTTCTGAGGAGGACGACTTCCGCCGTGTGGTCCGGGACGACGTGACCTGTTTCAT
CAGCGCGGTCCAGGACCAGGTGGTGCCGACAACACCTGGCCTGGGTGTGGGTGCGCGG
CCTGGACGAGCTGTACGCCGAGTGGTCCGAGGTGCTGTCCACGAACCTCCGGGACGCCTC
CGGGCCGGCCATGACCGAGATCGGCGAGCAGCCGTGGGGGCGGGAGTTGCCCCTGCGCGA
CCCGGCCGGCAACTGCGTGCACTTCTGCGCCGAGGAGCAGGACTGACACGTGCTAAAAC
TCATTTTTTAATTTAAAAGGATCTAGGTGAAGATCCTTTTTTGATAATCTCATGACCAAAAT
CCCTTAACGTGAGTTTTTCGTTCCACTGAGCGTCAGACCCCGTAGAAAAGATCAAAGGATC
TTCTTGAGATCCTTTTTTTCTGCGCGTAATCTGCTGCTTGCAAACAAAAAACCACCGCT
ACCAGCGGTGGTTTTGTTTGCCGGATCAAGAGCTACCAACTCTTTTTTCCGAAGGTAACCTGG
CTTCAGCAGAGCGCAGATACCAAATACTGTCCTTCTAGTGATAGCCGTAGTTAGGCCACCA
CTTCAAGAACTCTGTAGCACCGCCTACATACCTCGCTCTGCTAATCCTGTTACCAGTGGC
TGCTGCCAGTGGCGATAAGTCTGTCTTACCGGGTTGGACTCAAGACGATAGTTACCGGA
TAAGGCGCAGCGGTCCGGCTGAACGGGGGGTTCGTGCACACAGCCAGCTTGGAGCGAAC
GACCTACACCGAACTGAGATACCTACAGCGTGAGCTATGAGAAAGCGCCACGCTTCCCGA
AGGGAGAAAGGCGGACAGGTATCCGCTAAGCGGCAGGGTCGGAACAGGAGAGCGCACGAG
GGAGCTTCCAGGGGGAACGCCTGGTATCTTTATAGTCTGTGCGGGTTTCGCCACCTCTG
ACTTGAGCGTTCGATTTTTGTGATGCTCGTCAGGGGGGCGGAGCCTATGGAAAAACGCCAG
CAACGCGGCCTTTTTACGGTTCCTGGGCTTTTGCTGGCCTTTTGCTCACATGTTCTTTCC
TGCGTTATCCCTGATTCTGTGGATAACCGTATTACCGCCTTTGAGTGAGCTGATACCGC
TCGCCGAGCCGAACGACCGAGCGCAGCGAGTCAGTGAGCGAGGAAGCGGAAG

FIG. 18

SEQ ID NO: 11

GACTAGTACTGACGGACACACCGAAGCCCCGGCGGCAACCCTCAGCGGATGCCCCGGGGCTT
CACGTTTTCCAGGTCAGAAGCGGTTTTCGGGAGTAGTGCCCCAACTGGGGTAACCTTTGAG
TTCTCTCAGTTGGGGGCGTAGGGTCGCCGACATGACACAAGGGGTTGTGACCGGGGTGGACA
CGTACGCGGGTGCTTACGACCGTCAGTCGCGCGAGCGCGACTAGTACA

Fig. 19

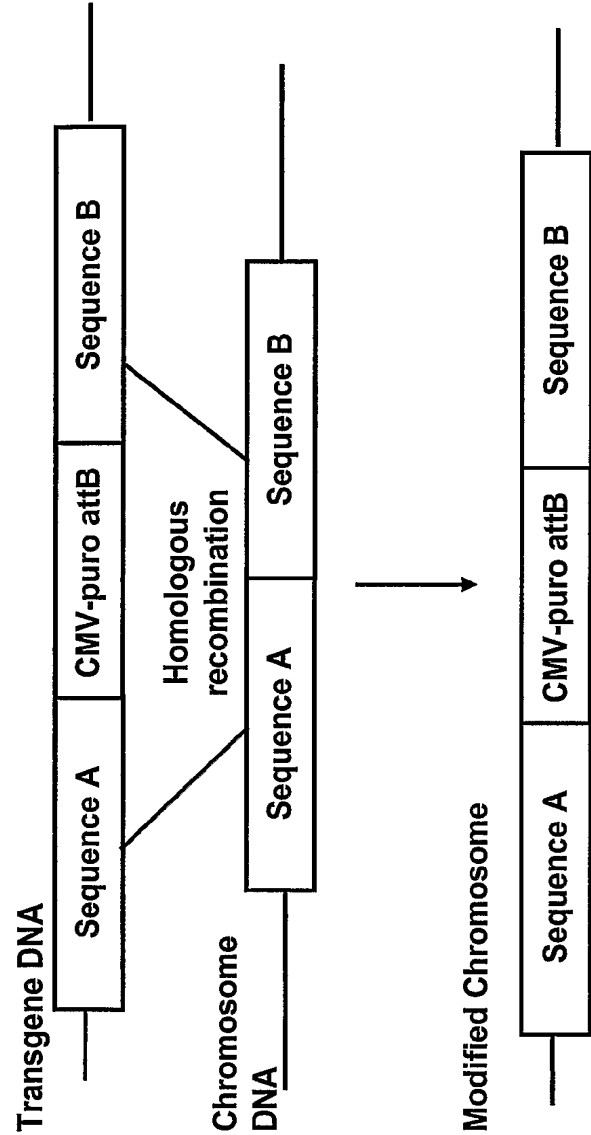


Fig. 20

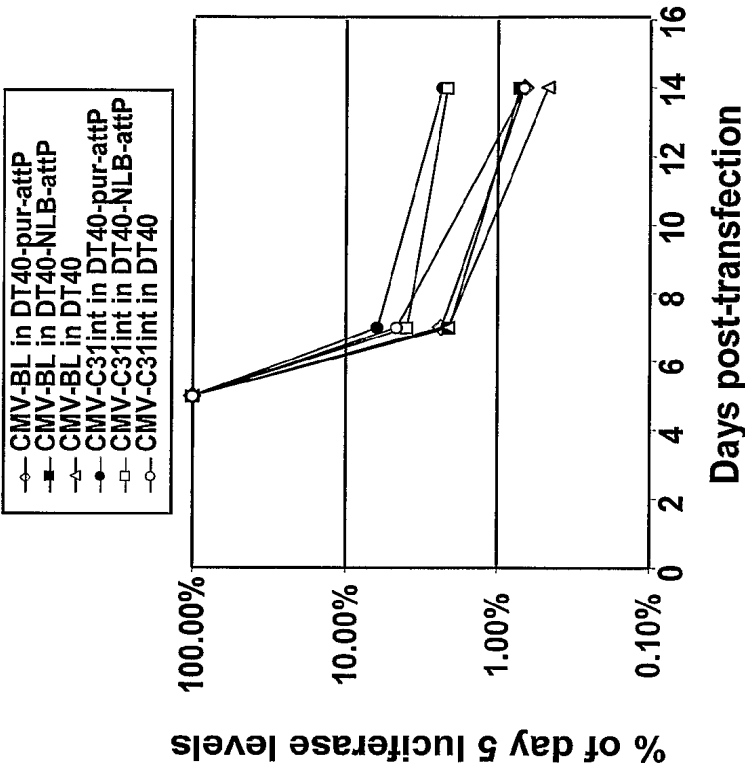


Fig. 21

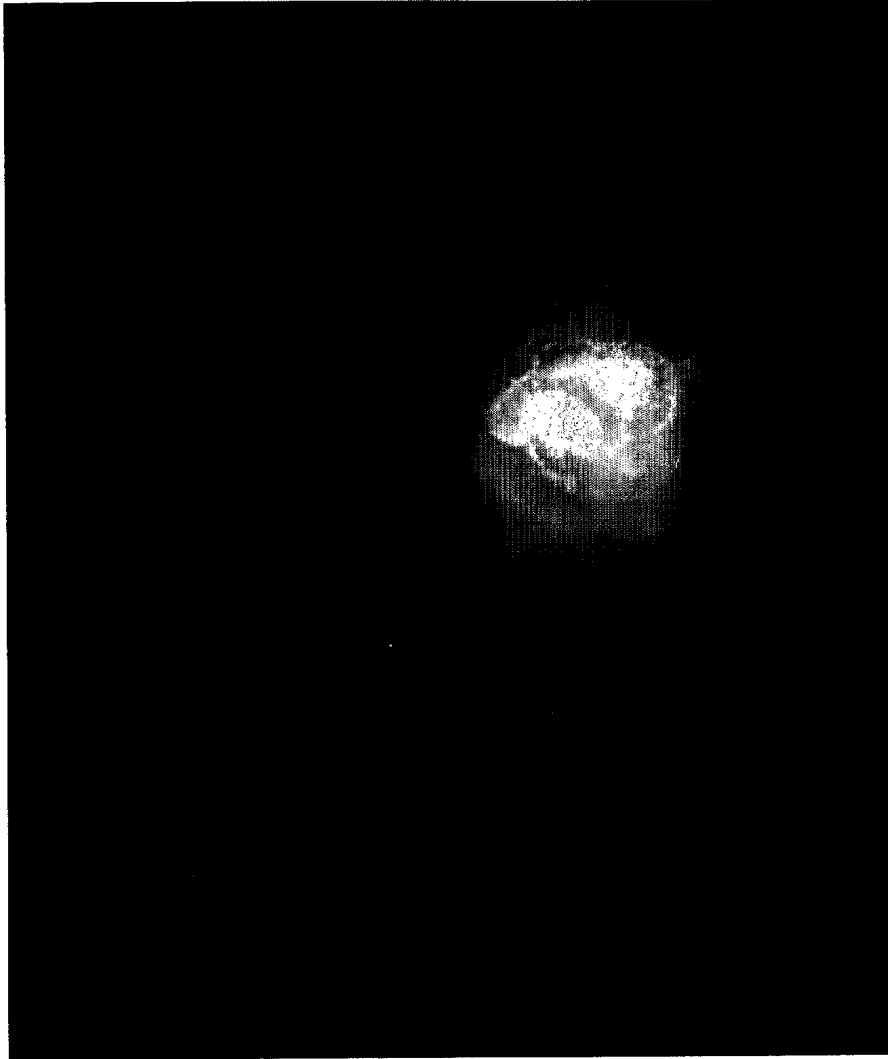


Fig. 22

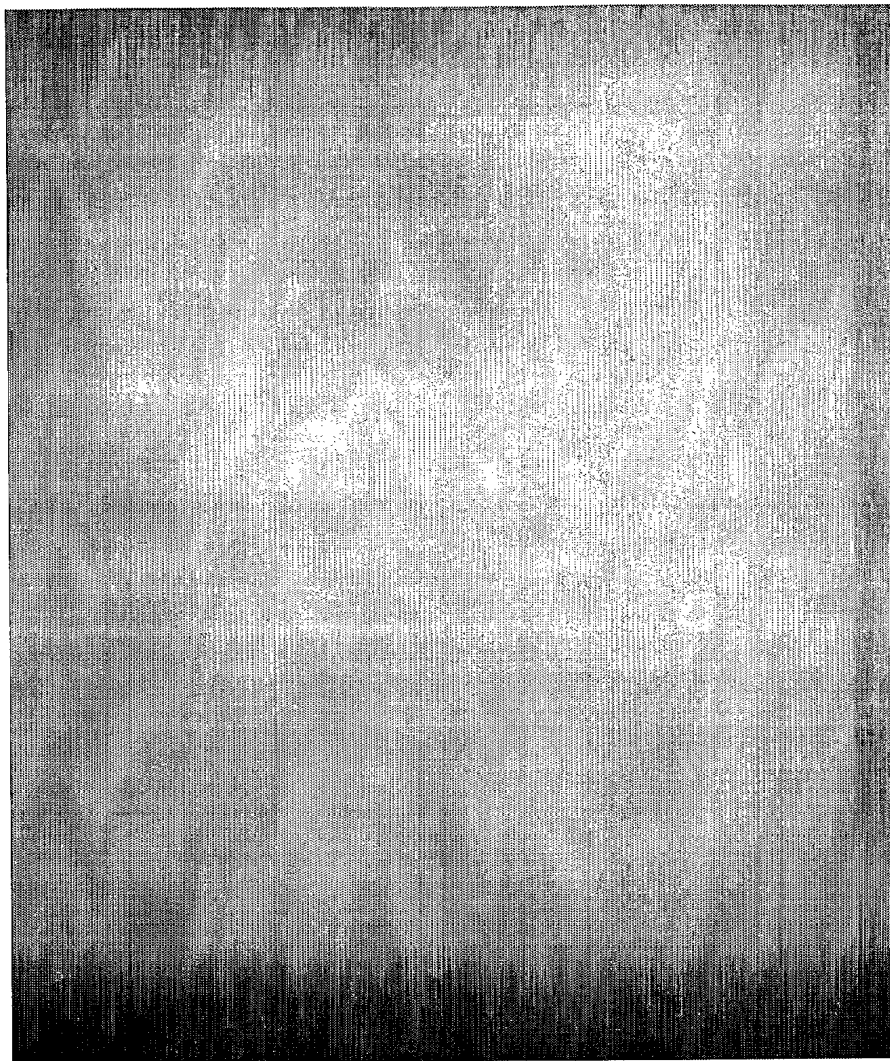


Fig. 23

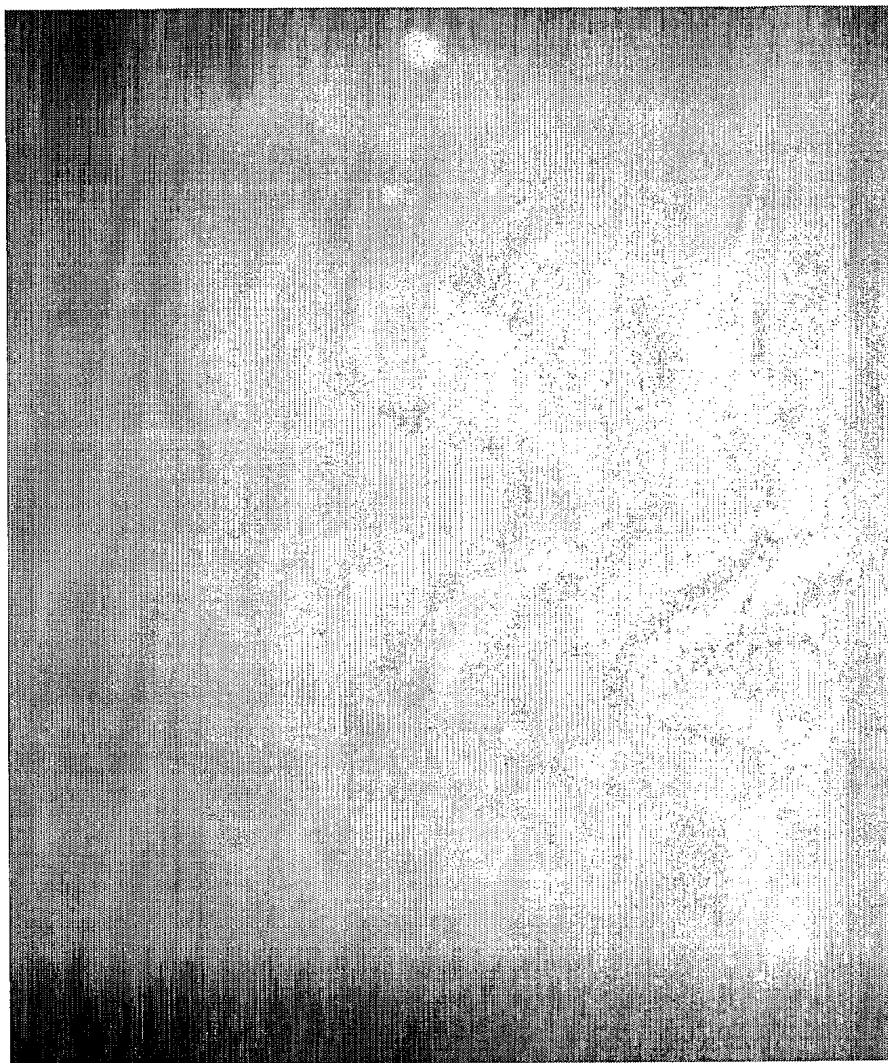


Fig. 24

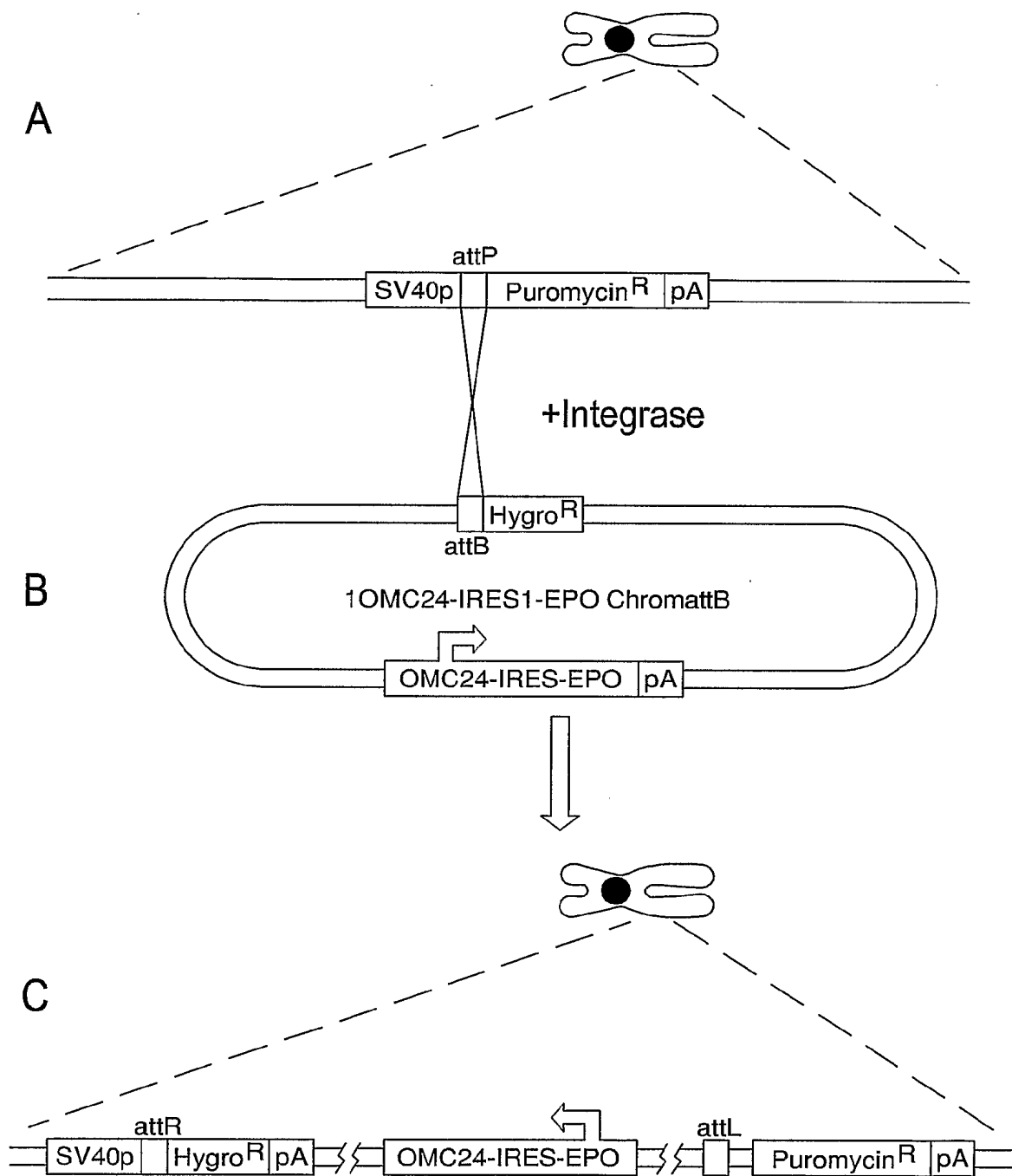


FIG. 25

025CIP4 PCT SEQ List.txt
SEQUENCE LISTING

<110> AviGenics, Inc
<120> MODIFICATION OF THE GENOME BY INTRODUCTION OF LARGE NUCLEIC ACID MOLECULES
<130> AVI-025CIP4PCT
<160> 13
<170> PatentIn version 3.2
<210> 1
<211> 6230
<212> DNA
<213> Artificial Sequence
<220>
<223> Plasmid pCMV-31int
<400> 1
cattcgccat tcaggctgcg caactgttgg gaagggcgat cgggtgcgggc ctcttcgcta 60
ttacgccagc caatacgcaa accgcctctc cccgcgcggt ggccgattca ttaatgcagg 120
atcgatccag acatgataag atacattgat gagtttggac aaaccacaac tagaatgcag 180
tgaaaaaat gctttatttg tgaaatttgg gatgctattg ctttatttgg aaccattata 240
agctgcaata aacaagttaa caacaacaat tgcattcatt ttatgtttca gggttcagggg 300
gaggtgtggg aggtttttta aagcaagtaa aacctctaca aatgtgggat ggctgattat 360
gatcatgaac agactgtgag gactgagggg cctgaaatga gccttgggac tgtgaatcta 420
aaatacaca acaattagaa tcactagctc ctgtgtataa tattttcata aatcatactc 480
agtaagcaaa actctcaagc agcaagcata tgcagctagt ttaacacatt atacacttaa 540
aaattttata tttaccttag agcttttaaat ctctgtagggt agtttgtcca attatgtcac 600
accacagaag taaggttcct tcacaaagat cccaagctag cttataatac gactcactat 660
agggagagag ctatgacgtc gcatgcacgc gtaagcttgg gcccctcgag ggatccgggt 720
gtctcgctac gccgctacgt cttccgtgcc gtcctgggcg tcgtcttcgt cgtcgtcggg 780
cggcggcttc gccacgtga tcgaagcgcg cttctcgatg ggcgttcctt gcccctgcc 840
cgtagtcgac ttcgtgacaa cgatcttgct tacgaagagc ccgacgaaca cgcgcttgct 900
gtctactgac gcgcgcccc accacgactt agggccgggt gggtcagcgt cggcgtcttc 960
ggggaaccat tgggtcaagg gaagcttcgg ggcttcggcg gcttcaagtt cggcaagccg 1020
ctcttcgcc ctttgcgtgc ggagcgtcag cgctgcctgt tgcttcggga agtgcttcct 1080
gccaacgggt ccgtcgtacg cgcctgccgc gcggtcttcg tacagctctt caagggcggt 1140
cagggcgctc gcgcgctccg caacaagggt cggcgttcg ccgctcttct caggcgcctc 1200
agtgcgctt ccgaagcgtc gggcggcttc ccacagaagc gccaacgtct cttcgtcgcc 1260

025CIP4 PCT SEQ List.txt

ttcggcgtgc ctgatcttgt tgaagatgcg ttccgcaacg aacttgctga gtgccgccat	1320
gctgacgttg cacgtgcctt cgtgctgccc aggtgctggac gggctcgacca cttccggcg	1380
acggcagcgg taagagtcct tgatcgattc ttccccgcgc ttcgaagtca tgacggcgcc	1440
acactcgag tacagcttgt ccatggcgga cagaatggct tgccccggg aaagcccctt	1500
gccgcgcccc ctgccgtcca accacgcctg aagctcatac cactcagcgg gctcgatgat	1560
cggtccgcaa tcaagctcga ccggccggag cgtgatcggg tcgcgctgaa tgcggtaacc	1620
ctcaatcttc gtggctggcg tgccgtccgg cttcttcttg tagatcacct cagcggcgaa	1680
gccccgaata cgcggtccc gaaggattcg cataacggtt gccgggtccc aggcgcttga	1740
agcggctctt ttcccaatcg tctcgccccg ggtcggcacg gcgtcagcgt ccatgcgctt	1800
acaaagcccc gtgatgctgc ccgggtgaat ggcggttga ctgcccggct tgaagggaag	1860
gtgtttgtgc gtcttgatct cacgccacca ccaccggatt acgtcgggct cgaactcgaa	1920
gggtccggta aggggagtg tgcagtgcgc aagcttggtg atgacgacat tgaccattcg	1980
gccgttgcg gtgatctcct tcgtctccga aacaagctcg aagccgtaag gcgccttccc	2040
gccgacgtac ccgccaatt cgcgctgaag gttcttcgtg tcgagaatct tcgccgactt	2100
cagcgaagat tctttgtgcg acgcgtcgag ccgcataatc aggtgaatca ggtccatgac	2160
gtttccctgc cggaagacgc cttcctgagt ggaaacaatc gtcacgcca gggcgagcaa	2220
ttccgagaca atcggaatcg cgtccatgac cttcaggcgc gagaagcgcg acacgtcata	2280
gacaatgatc atgttgagcc gccggcgcg gcattcggtc aggatgcgtt cgaactccgg	2340
gcgctccgcc gtcccgaacg ccgacgtgcc cggcgcttcg ctgaaatgcc cgacgaacct	2400
gaaccggccc ccgtcgcgct cgacttcgcg ctgaaggctg gccgccttgt cttcgttggc	2460
gctacgctgt gtcgctgggc ttgctgcgct cgaattctcg cgctcgcgcg actgacggtc	2520
gtaagcacc gcgtacgtgt ccacccgggt cacaaccct tgtgtcatgt cggcgaccct	2580
acgactagt agctcgtcga ccgggaatt ccggaccgggt acctgcaggc gtaccttcta	2640
tagtgtcacc taaatagctt ttgcaaaag cctaggctag agtccggagg ctggatcgggt	2700
cccgtgtct tctatggagg tcaaacagc gtggatggcg tctccaggcg atctgacgggt	2760
tcactaaac agctctgctt atatagacct ccacccgtac acgcctaccg cccatttgcg	2820
tcaatggggc ggagttgtta cgacattttg gaaagtcccc ttgattttgg tgccaaaaca	2880
aactcccatt gacgtcaatg ggggtggagac ttggaaatcc ccgtgagtca aaccgctatc	2940
cacgccatt gatgtactgc caaaaccgca tcaccatgggt aatagcgatg actaatacgt	3000
agatgtactg ccaagtagga aagtccata aggtcatgta ctgggcataa tgccaggcgg	3060
gccatttacc gtcattgacg tcaatagggg gcgtacttgg catatgatac acttgatgta	3120

025CIP4 PCT SEQ List.txt

ctgccaagtg ggcagttttac cgtaaataact ccaccatttg acgtcaatgg aaagtcacct	3180
ttggcggttac tatgggaaca tacgtcatta ttgacgtcaa tgggcggggg tcgttgggcg	3240
gtcagccagg cgggccatatt accgtaagtt atgtaacgac ctgcacgatg ctgtttcctg	3300
tgtgaaattg ttatccgctc acaattccac acattatacg agccggaagc tataaagtgt	3360
aaagcctggg gtgcctaatt agtgaaaggg cctcgtatac gcctatTTTT ataggTTaat	3420
gtcatgataa taatggTTTT ttagacgtca ggtggcactt ttcggggaaa tgtgcgcgga	3480
accctatatt gtttattttt ctaaatacat tcaaatatgt atccgctcat gagacaataa	3540
ccctgataaa tgcttcaata atattgaaaa acgcgcgaat tgcaagctct gcattaatga	3600
atcggccaac gcgcggggag aggcggtttg cgtattgggc gctcttccgc ttcctcgctc	3660
actgactcgc tgcgctcggt cgttcggctg cggcgagcgg tatcagctca ctcaaaggcg	3720
gtaatacggg tatccacaga atcaggggat aacgcaggaa agaacaatgt agcaaaaggc	3780
cagcaaaagg ccaggaaccg taaaaaggcc gcgttgctgg cgTTTTcca taggctccgc	3840
ccccctgacg agcatcaca aaatcgacgc tcaagtcaga ggtggcgaaa cccgacagga	3900
ctataaagat accaggcggt tccccctgga agctccctcg tgcgctctcc tgttccgacc	3960
ctgccgctta ccgatacct gtccgccttt ctcccttcgg gaagcgtggc gctttctcaa	4020
tgctcacgct gtaggtatct cagttcggtg taggtcggtt gctccaagct gggctgtgtg	4080
cacgaacccc ccgttcagcc cgaccgctgc gccttatccg gtaactatcg tcttgagtcc	4140
aaccggtaa gacacgactt atcgccactg gcagcagcca ctggtaacag gattagcaga	4200
gcgaggtatg taggcggtgc tacagagttc ttgaagtggg ggcctaacta cggctacact	4260
agaaggacag tatttggtat ctgcgctctg ctgaagccag ttaccttcgg aaaaagagtt	4320
ggtagctctt gatccggcaa acaaaccacc gctggtagcg gtggTTTTtt tgtttgcaag	4380
cagcagatta cgcgcagaaa aaaaggatct caagaagatc ctttgatctt ttctacgggg	4440
tctgacgctc agtggaaacga aaactcacgt taagggattt tggatcatgcc ataacttcgt	4500
atagcataca ttatacgaag ttatggcatg agattatcaa aaaggatctt cacctagatc	4560
cttttaaatt aaaaatgaag ttttaaata atctaaagta tatatgagta aacttggtct	4620
gacagttacc aatgcttaat cagtgaggca cctatctcag cgatctgtct atttcgttca	4680
tccatagttg cctgactccc cgtcgtgtag ataactacga tacgggaggg cttaccatct	4740
ggccccagtg ctgcaatgat accgcgagac ccacgctcac cggctccaga tttatcagca	4800
ataaaccagc cagccggaag ggccgagcgc agaagtggc ctgcaacttt atccgcctcc	4860
atccagtcta ttaattgttg ccgggaagct agagtaagta gtccgccagt taatagtttg	4920
cgcaacgttg ttgccattgc tacaggcatc gtggtgtcac gctcgtcgtt tggatggct	4980
tcattcagct ccggttccca acgatcaagg cgagttacat gatcccccat gttgtgcaaa	5040

025CIP4 PCT SEQ List.txt

```

aaagcggtta gctccttcgg tcctccgata gttgtcagaa gtaagttggc cgcagtgtta 5100
tcactcatgg ttatggcagc actgcataat tctcttactg tcatgccatc cgtaagatgc 5160
ttttctgtga ctggtgagta ctcaaccaag tcattctgag aatagtgtat gcggcgaccg 5220
agttgtcttt gcccggcgtc aatacgggat aataccgcgc cacatagcag aactttaaaa 5280
gtgctcatca ttggaaaacg ttcttcgggg cgaaaactct caaggatctt accgctgttg 5340
agatccagtt cgatgtaacc cactcgtgca cccaactgat cttcagcatc ttttactttc 5400
accagcgttt ctgggtgagc aaaaacagga aggcaaatg ccgcaaaaaa gggaataagg 5460
gcgacacgga aatgttgaat actcatactc ttcctttttc aatattattg aagcatttat 5520
cagggttatt gtctcatgcc aggggtgggc acacatattt gataccagcg atccctacac 5580
agcacataat tcaatgcgac ttccctctat cgcacatctt agacctttat tctccctcca 5640
gcacacatcg aagctgccga gcaagccgtt ctcaccagtc caagacctgg catgagcgga 5700
tacatatttg aatgtattta gaaaaataaa caaatagggg ttccgcgcac atttccccga 5760
aaagtgccac ctgaaattgt aaacgttaat attttgttaa aattcgcgtt aaatTTTTgt 5820
taaatacagc cattttttta ccaataggcc gaaatcggca aaatccctta taaatcaaaa 5880
gaatagaccg agatagggtt gagtgttggt ccagtttgga acaagagtcc actattaaag 5940
aacgtggact ccaacgtcaa agggcgaaaa accgtctatc agggcgatgg cccactacgt 6000
gaaccatcac cctaatacag ttttttgggg tcgaggtgcc gtaaagcact aaatcggaac 6060
cctaaaggga gccccgatt tagagcttga cggggaaagc cggcgaacgt ggcgagaaag 6120
gaagggaaga aagcgaaagg agcgggcgct agggcgctgg caagtgtagc ggtcacgctg 6180
cgcgtaacca ccacaccgc cgcgcttaat gcgcgcgtac agggcgcgctc 6230

```

<210> 2
 <211> 5982
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> Plasmid pCMV-luc-attB

```

<400> 2
ctctatcgat aggtaccgag ctcttacgcg tgctagccct cgagcaggat ctatacattg 60
aatcaatatt ggcaattagc catattagtc attgggtata tagcataaat caatattggc 120
tattggccat tgcatacgtt gtatctatat cataatatgt acatttatat tggctcatgt 180
ccaatatgac cgccatgttg acattgatta ttgactagtt attaatagta atcaattacg 240
gggtcattag ttcatagccc atatatggag ttccgcgtta cataacttac ggtaaattggc 300
ccgcctggct gaccgccccaa cgacccccgc ccattgacgt caataatgac gtatgttccc 360

```

025CIP4 PCT SEQ List.txt

atagtaacgc	caatagggac	tttccattga	cgtcaatggg	tgagagtattt	acggtaaact	420
gcccacttgg	cagtacatca	agtgtatcat	atgccaagtc	cgccccctat	tgacgtcaat	480
gacggtaaatt	ggcccgccctg	gcattatgcc	cagtacatga	ccttacggga	ctttcctact	540
tggcagtaca	tctacgtatt	agtcacgcgt	attaccatgg	tgatgcgggtt	ttggcagtac	600
atcaatgggc	gtggatagcg	gtttgactca	cggggatttc	caagtctcca	ccccattgac	660
gtcaatggga	gtttgttttg	gcacaaaaat	caacgggact	ttccaaaatg	tcgtaacaac	720
tccgccccat	tgacgcaaat	gggcggtagg	cgtgtacggg	gggaggtcta	tataagcaga	780
gctcgtttag	tgaaccgtca	gatgccttgg	agacgccatc	cacgctgttt	tgacctccat	840
agaagacacc	gggaccgatc	cagcctcccc	tcgaagctcg	actctagggg	ctcgagatct	900
gcgatctaag	taagcttggc	attccgggtac	tgttggtaaa	gccaccatgg	aagacgccaa	960
aaacataaag	aaaggcccg	cgccatttcta	tccgctggaa	gatggaaccg	ctggagagca	1020
actgcataag	gctatgaaga	gatacgccct	ggttcctgga	acaattgctt	ttacagatgc	1080
acatatcgag	gtggacatca	cttacgctga	gtacttcgaa	atgtccgttc	ggttggcaga	1140
agctatgaaa	cgatatgggc	tgaatacaaa	tcacagaatc	gtcgtatgca	gtgaaaactc	1200
tcttcaattc	tttatgccgg	tgttgggcgc	gttattttatc	ggagttgcag	ttgcgcccgc	1260
gaacgacatt	tataatgaac	gtgaattgct	caacagtatg	ggcattttcgc	agcctaccgt	1320
gggtgttcgtt	tccaaaaagg	ggttgcaaaa	aattttgaac	gtgcaaaaaa	agctcccaat	1380
catccaaaaa	attattatca	tggaattctaa	aacggattac	cagggatttc	agtcgatgta	1440
cacgttcgtc	acatctcatc	tacctcccg	ttttaatgaa	tacgattttg	tgccagagtc	1500
cttcgatagg	gacaagacaa	ttgcactgat	catgaactcc	tctggatcta	ctggtctgcc	1560
taaagggtgtc	gctctgcctc	atagaactgc	ctgcgtgaga	ttctcgcgatg	ccagagatcc	1620
tattttttggc	aatcaaatca	ttccggatac	tgcgatttta	agtgttggtc	cattccatca	1680
cggtttttggg	atgtttacta	cactcggtata	tttgatatgt	ggattttcgag	tcgtcttaat	1740
gtatagatatt	gaagaagagc	tgtttctgag	gagccttcag	gattacaaga	ttcaaagtgc	1800
gctgctgggtg	ccaaccctat	tctccttctt	cgccaaaagc	actctgattg	acaaatacga	1860
tttatctaatt	ttacacgaaa	ttgcttctg	tggcgctccc	ctctctaagg	aagtcggggg	1920
agcggttgcc	aagaggttcc	atctgccagg	tatcaggcaa	ggatatgggc	tcactgagac	1980
tacatcagct	attctgatta	cacccgaggg	ggatgataaa	ccgggcgcgg	tcggtaaagt	2040
tgttccattt	tttgaagcga	aggttggtga	tctggatacc	gggaaaacgc	tgggcgttaa	2100
tcaaagaggc	gaactgtgtg	tgagagggtcc	tatgattatg	tccggttatg	taaacaatcc	2160
ggaagcgacc	aacgccttga	ttgacaagga	tggtatggcta	cattctggag	acatagctta	2220
ctgggacgaa	gacgaacact	tcttcatcgt	tgaccgcctg	aagtctctga	ttaagtacaa	2280

025CIP4 PCT SEQ List.txt

aggctatcag	gtggctccc	ctgaattgga	atccatcttg	ctccaacacc	ccaacatctt	2340
cgacgcaggt	gtcgcaggtc	ttcccgcagc	tgacgccggt	gaacttccc	ccgccgttgt	2400
tgttttgag	cacggaaaga	cgatgacgga	aaaagagatc	gtggattacg	tcgccagtca	2460
agtaacaacc	gcgaaaaagt	tgcgcggagg	agttgtgttt	gtggacgaag	taccgaaagg	2520
tcttaccgga	aaactcgacg	caagaaaaat	cagagagatc	ctcataaagg	ccaagaaggg	2580
cggaaagatc	gccgtgtaat	tctagagtcg	ggcgggccgg	ccgcttcgag	cagacatgat	2640
aagatacatt	gatgagtttg	gacaaaccac	aactagaatg	cagtgaaaaa	aatgctttat	2700
ttgtgaaatt	tgtgatgcta	ttgctttatt	tgttaaccatt	ataagctgca	ataaacaagt	2760
taacaacaac	aattgcattc	attttatgtt	tcaggttcag	ggggaggtgt	gggaggtttt	2820
ttaaagcaag	taaaacctct	acaaatgtgg	taaaatcgat	aaggatcaat	tcggcttcag	2880
gtaccgtcga	cgatgtaggt	cacggctctc	aagccgcggt	gcgggtgcc	gggcgtgccc	2940
ttgggctccc	cgggcgcgta	ctccacctca	cccatctggt	ccatcatgat	gaacgggtcg	3000
aggtggcggg	agttgatccc	ggcgaacgcg	cggcgcaccg	ggaagccctc	gccctcgaaa	3060
ccgctgggcg	cgggtggcac	ggtgagcacg	ggacgtgcga	cggcgtcggc	gggtgcggat	3120
acgcggggca	gcgtcagcgg	gttctcgacg	gtcacggcgg	gcatgtcgac	agccgaattg	3180
atccgtcgac	cgatgccctt	gagagccttc	aaccagtcga	gtccttccg	gtgggcgcgg	3240
ggcatgacta	tcgtcgccgc	acttatgact	gtcttcttta	tcatgcaact	cgtaggacag	3300
gtgccggcag	cgctcttccg	cttcctcgct	cactgactcg	ctgcgctcgg	tcgttcggct	3360
gcggcgagcg	gtatcagctc	actcaaaggc	ggtaatacgg	ttatccacag	aatcagggga	3420
taacgcagga	aagaacatgt	gagcaaaaag	ccagcaaaaag	gccaggaacc	gtaaaaaggc	3480
cgcgttgctg	gcgtttttcc	ataggctccg	ccccctgac	gagcatcaca	aaaatcgacg	3540
ctcaagtcag	aggtggcgaa	acccgacagg	actataaaga	taccaggcgt	ttccccctgg	3600
aagctccctc	gtgcgctctc	ctgttccgac	cctgccgctt	accggatacc	tgtccgcctt	3660
tctcccttcg	ggaagcgtgg	cgctttctca	atgctcacgc	tgtaggtatc	tcagttcggt	3720
gtaggtcggt	cgctccaagc	tgggctgtgt	gcacgaacct	cccgttcagc	ccgaccgctg	3780
cgccttatcc	ggtaactatc	gtcttgagtc	caaccgggta	agacacgact	tatcgccact	3840
ggcagcagcc	actggtaaca	ggattagcag	agcgaggtat	gtaggcggtg	ctacagagtt	3900
cttgaagtgg	tggcctaact	acggctacac	tagaaggaca	gtatttggtg	tctgcgctct	3960
gctgaagcca	gttaccttcg	gaaaaagagt	tggtagctct	tgatccggca	aacaaaccac	4020
cgtggtagc	ggtgggtttt	ttgtttgcaa	gcagcagatt	acgcgcagaa	aaaaaggatc	4080
tcaagaagat	cctttgatct	tttctacggg	gtctgacgct	cagtggaacg	aaaactcacg	4140

025CIP4 PCT SEQ List.txt

ttaagggatt ttgggtcatga gattatcaaaa aaggatcttc acctagatcc ttttaaatta	4200
aaaatgaagt tttaaatcaa tctaaagtat atatgagtaa acttggtctg acagttacca	4260
atgcttaatc agtgaggcac ctatctcagc gatctgtcta tttcgttcat ccatagttgc	4320
ctgactcccc gtcgtgtaga taactacgat acgggagggc ttaccatctg gccccagtgc	4380
tgcaatgata ccgcgagacc cacgctcacc ggctccagat ttatcagcaa taaaccagcc	4440
agccggaagg gccgagcgca gaagtgggtcc tgcaacttta tccgcctcca tccagtctat	4500
taattgtttgc cggaagcta gagtaagtag ttcgccagtt aatagtttgc gcaacgttgt	4560
tgccattgct acaggcatcg tgggtgtcacg ctcgctgttt ggtatggctt cattcagctc	4620
cggttcccaa cgatcaaggc gagttacatg atcccccagc ttgtgcaaaa aagcggttag	4680
ctccttcggt cctccgatcg ttgtcagaag taagttggcc gcagtgttat cactcatggt	4740
tatggcagca ctgcataatt ctcttactgt catgccatcc gtaagatgct tttctgtgac	4800
tgggtgagtac tcaaccaagt cattctgaga atagtgtatg cggcgaccga gttgctcttg	4860
cccggcgtca atacgggata ataccgcgcc acatagcaga actttaaaag tgctcatcat	4920
tggaaaacgt tcttcggggc gaaaactctc aaggatctta ccgctgttga gatccagttc	4980
gatgtaaccc actcgtgcac ccaactgatc ttcagcatct tttactttca ccagcgtttc	5040
tgggtgagca aaaacaggaa ggcaaaaatgc cgcaaaaaag ggaataaggg cgacacggaa	5100
atgttgaata ctcatactct tcctttttca atattattga agcatttatc agggttattg	5160
tctcatgagc ggatacatat ttgaatgtat ttagaaaaat aaacaaatag gggttccgcg	5220
cacatttccc cgaaaagtgc cacctgacgc gccctgtagc ggcgcattaa gcgcggcggg	5280
tgtggtggtt acgcgcagcg tgaccgctac acttgccagc gccctagcgc ccgctccttt	5340
cgctttcttc ccttcctttc tcgccacgtt cgccggcttt ccccgtaag ctctaaatcg	5400
ggggctccct ttagggttcc gatttagtgc tttacggcac ctcgaccca aaaaacttga	5460
ttaggggtgat ggttcacgta gtggggccatc gccctgatag acggtttttc gcccttgac	5520
gttggagtcc acgttcttta atagtggact cttgttccaa actggaacaa cactcaaccc	5580
tatctcggtc tattcttttg atttataagg gattttgccg atttcggcct attgggttaa	5640
aatgagctg atttaacaaa aatttaacgc gaattttaac aaaatattaa cgtttacaat	5700
ttcccattcg ccattcaggc tgcgcaactg ttgggaaggg cgatcgggtgc gggcctcttc	5760
gctattacgc cagcccaagc taccatgata agtaagtaat attaaggtac gggaggtact	5820
tggagcggcc gcaataaaat atctttatct tcattacatc tgtgtgttgg ttttttgtgt	5880
gaatcgatag tactaacata cgctctccat caaaacaaaa cgaaacaaaa caaactagca	5940
aataggtctg tccccagtgc aagtgcaggt gccagaacat tt	5982

025CIP4 PCT SEQ List.txt

<210> 3
 <211> 5924
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> Plasmid pCMV-luc-attP

<400> 3
 ctctatcgat aggtaccgag ctcttacgcg tgctagccct cgagcaggat ctatacattg 60
 aatcaatatt ggcaattagc catattagtc attgggtata tagcataaat caatattggc 120
 tattggccat tgcatacggt gtatctatat cataatatgt acattttatat tggctcatgt 180
 ccaatatgac cgccatgttg acattgatta ttgactagtt attaatagta atcaattacg 240
 gggtcattag ttcatagccc atatatggag ttccgcgtta cataacttac ggtaaattggc 300
 ccgcctggct gaccgcccga cgacccccgc ccattgacgt caataatgac gtatgttccc 360
 atagtaacgc caatagggac tttccattga cgtcaatggg tggagtattt acggtaaact 420
 gccacttg cagtacatca agtgtatcat atgccaagtc cgccccctat tgacgtcaat 480
 gacggtaaatt ggccgcctg gcattatgcc cagtacatga ccttacggga ctttcctact 540
 tggcagtaca tctacgtatt agtcatcgct attaccatgg tgatgcgggt ttggcagtac 600
 atcaatgggc gtggatagcg gtttgactca cggggatttc caagtctcca cccattgac 660
 gtcaatggga gtttggtttg gcacaaaat caacgggact ttcaaaaatg tcgtaacaac 720
 tccgccccat tgacgcaaat gggcggtagg cgtgtacggg gggagggtcta tataagcaga 780
 gctcgttttag tgaaccgtca gatgcgctgg agacgccatc cagctgtttt tgacctccat 840
 agaagacacc gggaccgatc cagcctcccc tcgaagctcg actctagggg ctcgagatct 900
 gcgatctaag taagcttggc attccggtac tgttggtaaa gccaccatgg aagacgcca 960
 aaacataaag aaaggcccg cgccattcta tccgctggaa gatggaaccg ctggagagca 1020
 actgcataag gctatgaaga gatacgccct gggttcctgga acaattgctt ttacagatgc 1080
 acatatcgag gtggacatca cttacgctga gtacttcgaa atgtccgttc gggtggcaga 1140
 agctatgaaa cgatatgggc tgaatacaaa tcacagaatc gtcgtatgca gtgaaaactc 1200
 tcttcaattc tttatgccgg tgttgggcgc gttatttatc ggagttgcag ttgcgccccg 1260
 gaacgacatt tataatgaac gtgaattgct caacagtatg ggcatttcgc agcctaccgt 1320
 ggtgttcgtt tccaaaaagg ggttgcaaaa aattttgaac gtgcaaaaaa agctcccaat 1380
 catcaaaaaa attattatca tggattctaa aacggattac cagggtattc agtcgatgta 1440
 cacgttcgtc acatctcatc tacctcccgg ttttaatgaa tacgattttg tgccagagtc 1500
 cttcgatagg gacaagacaa ttgcactgat catgaactcc tctggatcta ctggtctgcc 1560
 taaagggtgtc gctctgcctc atagaactgc ctgcgtgaga ttctcgcgat ccagagatcc 1620

025CIP4 PCT SEQ List.txt

tatTTTTTggc	aatcaaataca	ttccggatac	tgcgatttta	agtgttggtc	cattccatca	1680
cggtTTTtgga	atgtttacta	cactcggata	tttgatatgt	ggatttccgag	tcgtcttaata	1740
gtatagatttt	gaagaagagc	tgTTTTctgag	gagccttcag	gattacaaga	ttcaaagtgc	1800
gctgctggtg	ccaaccctat	tctccttctt	cgccaaaagc	actctgattg	acaaatacga	1860
tttatctaata	ttacacgaaa	ttgcttctg	tggcgctccc	ctctctaagg	aagtcggggga	1920
agcggttgcc	aagaggttcc	atctgccagg	tatcaggcaa	ggatatgggc	tcactgagac	1980
tacatcagct	attctgatta	cacccgaggg	ggatgataaa	cggggcgcg	tcggttaaagt	2040
tgttccatttt	tttgaagcga	aggttgtgga	tctggatacc	gggaaaacgc	tgggcgttaa	2100
tcaaagaggc	gaactgtgtg	tgagaggtcc	tatgattatg	tccggttatg	taaacaatcc	2160
ggaagcgacc	aacgccttga	ttgacaagga	tggatggcta	cattctggag	acatagctta	2220
ctgggacgaa	gacgaacact	tcttcatcgt	tgaccgcctg	aagtctctga	ttaagtacaa	2280
aggctatcag	gtggctcccg	ctgaattgga	atccatcttg	ctccaacacc	ccaacatctt	2340
cgacgcaggt	gtcgcaggtc	ttcccgacga	tgacgccggt	gaacttcccg	ccgccgttgt	2400
tgTTTTggag	cacggaaaga	cgatgacgga	aaaagagatc	gtggattacg	tcgccagtca	2460
agtaacaacc	gcgaaaaagt	tgcgcgagg	agttgtgttt	gtggacgaag	taccgaaagg	2520
tcttaccgga	aaactcgacg	caagaaaaat	cagagagatc	ctcataaagg	ccaagaagg	2580
cggaaagatc	gccgtgtaata	tctagagtcg	ggcgggcccg	ccgcttcgag	cagacatgat	2640
aagatacatt	gatgagtttg	gacaaaccac	aactagaatg	cagtgaaaaa	aatgctttat	2700
ttgtgaaatt	tgtgatgcta	ttgctttatt	tgtaaccatt	ataagctgca	ataaacaagt	2760
taacaacaac	aattgcattc	atTTTtatgtt	tcaggttcag	ggggagggtgt	gggaggTTTT	2820
ttaaagcaag	taaaacctct	acaaatgtgg	taaaatcgat	aaggatcaat	tcggcttcga	2880
ctagtactga	cggacacacc	gaagccccgg	cggcaaccct	cagcggatgc	cccggggctt	2940
cacgttttcc	caggtcagaa	gcggttttctg	ggagtagtgc	cccaactggg	gtaacctttg	3000
agtTctctca	gttggggggcg	tagggctcgcc	gacatgacac	aaggggttgt	gaccgggggtg	3060
gacacgtacg	cgggtgctta	cgaccgtcag	tcgcgcgagc	gcgactagta	caagccgaat	3120
tgatccgtcg	accgatgccc	ttgagagcct	tcaaccagat	cagctccttc	cgggtgggcgc	3180
ggggcatgac	tatcgtcgcc	gcacttatga	ctgtcttctt	tatcatgcaa	ctcgtaggac	3240
aggtgccggc	agcgctcttc	cgcttcctcg	ctcactgact	cgctgcgctc	ggtcgttcgg	3300
ctgcggcgag	cggtatcagc	tcactcaaag	gcggtaatac	ggttatccac	agaatcaggg	3360
gataacgcag	gaaagaacat	gtgagcaaaa	ggccagcaaa	aggccaggaa	ccgtaaaaag	3420
gccgcgttgc	tggcgTTTTT	ccataggctc	cgccccctg	acgagcatca	caaaaatcga	3480
cgctcaagtc	agagggtggcg	aaacccgaca	ggactataaa	gataccaggc	gtttccccct	3540

025CIP4 PCT SEQ List.txt

ggaagctccc	tcgtgcgctc	tcctgttccg	accctgccgc	ttaccggata	cctgtccgcc	3600
tttctccctt	cgggaagcgt	ggcgctttct	caatgctcac	gctgtaggta	tctcagttcg	3660
gtgtaggtcg	ttcgctccaa	gctgggctgt	gtgcacgaac	cccccgttca	gccccgaccgc	3720
tgcgcccttat	ccggttaacta	tcgtcttgag	tccaacccgg	taagacacga	cttatcgcca	3780
ctggcagcag	ccactggtaa	caggattagc	agagcgaggt	atgtaggcgg	tgctacagag	3840
ttcttgaaagt	ggtggcctaa	ctacggctac	actagaagga	cagtattttgg	tatctgcgct	3900
ctgctgaagc	cagttacctt	cggaaaaaga	gttggttagct	cttgatccgg	caaacaaacc	3960
accgctggta	gcggtggttt	ttttgtttgc	aagcagcaga	ttacgcgcag	aaaaaaagga	4020
tctcaagaag	atcctttgat	cttttctacg	gggtctgacg	ctcagtggaa	cgaaaactca	4080
cgtaagggga	ttttggtcat	gagattatca	aaaaggatct	tcacctagat	ccttttaaata	4140
taaaaatgaa	gttttaaatac	aatctaaagt	atatatgagt	aaacttggtc	tgacagttac	4200
caatgcttaa	tcagtgaggc	acctatctca	gcgatctgtc	tatttcgttc	atccatagtt	4260
gcctgactcc	ccgtcgtgta	gataactacg	atacgggagg	gcttaccatc	tggccccagt	4320
gctgcaatga	taccgcgaga	cccacgctca	ccggctccag	atttatcagc	aataaaccag	4380
ccagccggaa	gggccgagcg	cagaagtggg	cctgcaactt	tatccgcctc	catccagtct	4440
attaattggt	gccgggaagc	tagagtaagt	agttcgccag	ttaatagttt	gcgcaacggt	4500
gttgccattg	ctacaggcat	cgtgggtgtca	cgctcgtcgt	ttggatatggc	ttcattcagc	4560
tccggttccc	aacgatcaag	gcgagttaca	tgatcccca	tgttgtgcaa	aaaagcggtt	4620
agctccttcg	gtcctccgat	cgttgctcaga	agtaagttgg	ccgcagtgtt	atcactcatg	4680
gttatggcag	cactgcataa	ttctcttact	gtcatgccat	ccgtaagatg	cttttctgtg	4740
actggtgagt	actcaaccaa	gtcattctga	gaatagtgtg	tgcggcgacc	gagttgctct	4800
tgccccgcgt	caatacggga	taataccgcg	ccacatagca	gaactttaaa	agtgtcatc	4860
attggaaaac	gttcttcggg	gcgaaaactc	tcaaggatct	taccgctgtt	gagatccagt	4920
tcgatgtaac	ccactcgtgc	acccaactga	tcttcagcat	cttttacttt	caccagcggt	4980
tctgggtgag	caaaaacagg	aaggcaaaat	gccgcaaaaa	aggggaataag	ggcgacacgg	5040
aaatgttgaa	tactcatact	cttccttttt	caatattatt	gaagcattta	tcagggttat	5100
tgtctcatga	gcggatacat	atttgaatgt	atttagaaaa	ataaacaat	aggggttccg	5160
cgcacatttc	cccgaaaagt	gccacctgac	gcgccctgta	gcggcgcat	aagcgcgcg	5220
ggtgtggtgg	ttacgcgcag	cgtgaccgct	acacttgcca	gcgccctagc	gcccgtcct	5280
ttcgctttct	tcccttcctt	tctcgccacg	ttcgccggct	ttccccgtca	agctctaaat	5340
cgggggctcc	ctttaggggt	ccgatttagt	gctttacggc	acctcgaccc	caaaaaactt	5400

025CIP4 PCT SEQ List.txt

gattaggggtg atgggttcacg tagtggggcca tcgccctgat agacgggtttt tcgccctttg	5460
acgttggagt ccacgtttctt taatagtggga ctcttggtcc aaactggaac aacactcaac	5520
cctatctcgg tctattcttt tgatttataa gggattttgc cgatttcggc ctattggtta	5580
aaaaatgagc tgatttaaca aaaatttaac gcgaatttta acaaaatatt aacgtttaca	5640
atttcccatt cgccattcag gctgcgcaac tgttgggaag ggcgatcggg gcgggcctct	5700
tcgctattac gccagcccaa gctaccatga taagtaagta atattaaggt acgggaggtta	5760
cttggagcgg ccgcaataaa atatctttat ttccattaca tctgtgtgtt ggttttttgt	5820
gtgaatcgat agtactaaca tacgctctcc atcaaaacaa aacgaaacaa aacaaactag	5880
caaaataggc tgtccccagt gcaagtgcag gtgccagaac attt	5924

<210> 4

<211> 5101

<212> DNA

<213> Artificial Sequence

<220>

<223> Plasmid pCMV-pur-attB

<400> 4

ctagagtcgg ggccggccggc cgcttcgagc agacatgata agatacattg atgagtttgg	60
acaaaccaca actagaatgc agtgaaaaaa atgctttatt tgtgaaattt gtgatgctat	120
tgctttatatt gtaaccatta taagctgcaa taaacaagtt aacaacaaca attgcattca	180
ttttatgttt cagggttcagg gggagggtgtg ggagggttttt taaagcaagt aaaacctcta	240
caaatgtggg aaaatcgata aggatcaatt cggcttcagg taccgtcgac gatgtaggtc	300
acgggtctcga agccgcgggtg cgggtgccag ggcgtgccct tgggctcccc gggcgcgtac	360
tccacctcac ccatctggtc catcatgatg aacgggtcga ggtggcggtta gttgatcccc	420
gcgaacgcgc ggccgcaccgg gaagccctcg ccctcgaaac cgctgggcgc ggtgggtcacg	480
gtgagcacgg gacgtgcgac ggcgtcggcg ggtgcggata cgcggggcag cgtcagcggg	540
ttctcgacgg tcacggcggg catgtcgaca gccgaattga tccgtcgacc gatgcccttg	600
agagccttca acccagtcag ctcttccgg tgggcgcggg gcatgactat cgtcgccgca	660
cttatgactg tcttctttat catgcaactc gtaggacagg tgccggcagc gctcttccgc	720
ttcctcgctc actgactcgc tgcgtcgggt cgcttcggctg cggcgagcgg tatcagctca	780
ctcaaaggcg gtaatacggg tatccacaga atcaggggat aacgcaggaa agaacatgtg	840
agcaaaaggc cagcaaaagg ccaggaaccg taaaaaggcc gcgttgctgg cgtttttcca	900
taggctccgc cccctgacg agcatcaca aaatcgacgc tcaagtcaga ggtggcgaaa	960
cccgcagga ctataaagat accaggcgtt tccccctgga agctccctcg tgcgctctcc	1020
tgttccgacc ctgccgctta ccggatacct gtccgccttt ctcccttcgg gaagcgtggc	1080

025CIP4 PCT SEQ List.txt

gcttttctcaa tgctcacgct gtaggtatct cagttcggtg taggtcgttc gctccaagct	1140
gggctgtgtg cacgaacccc ccgttcagcc cgaccgctgc gccttatccg gtaactatcg	1200
tcttgagtcc aacccggtaa gacacgactt atcgccactg gcagcagcca ctggtaacag	1260
gattagcaga gcgaggtatg taggcggtgc tacagagttc ttgaagtggg ggcctaacta	1320
cggctacact agaaggacag tatttgggtat ctgcgctctg ctgaagccag ttaccttcgg	1380
aaaaagagtt ggtagctctt gatccggcaa acaaacacc gctggtagcg gtgggtttttt	1440
tgtttgcaag cagcagatta cgcgcagaaa aaaaggatct caagaagatc ctttgatctt	1500
ttctacgggg tctgacgctc agtggaacga aaactcacgt taagggattt tgggtcatgag	1560
attatcaaaa aggatcttca cctagatcct tttaaattaa aaatgaagtt ttaaataaat	1620
ctaaagtata tatgagtaaa cttgggtctga cagttacca tgcttaatca gtgaggcacc	1680
tatctcagcg atctgtctat ttcggttcac catagttgcc tgactccccg tcgtgtagat	1740
aactacgata cgggagggct taccatctgg cccagtgct gcaatgatac cgcgagaccc	1800
acgctcaccg gctccagatt tatcagcaat aaaccagcca gccggaaggg ccgagcgag	1860
aagtggctct gcaactttat ccgcctccat ccagtcctatt aattgttgcc ggaagctag	1920
agtaagtagt tcgccagtta atagtttgcg caacgttggt gccattgcta caggcatcgt	1980
gggtgtcacgc tcgtcgtttg gtatggcttc attcagctcc ggttcccaac gatcaaggcg	2040
agttacatga tccccatgt tgtgcaaaaa agcgggttagc tccttcggtc ctccgatcgt	2100
tgtcagaagt aagttggccg cagtgttatc actcatgggt atggcagcac tgcataattc	2160
tcttactgtc atgccatccg taagatgctt ttctgtgact ggtgagtact caaccaagtc	2220
attctgagaa tagtgtatgc ggcgaccgag ttgctcttgc ccggcgtaa tacgggataa	2280
taccgcgcca catagcagaa ctttaaaagt gctcatcatt ggaaaacgtt cttcggggcg	2340
aaaactctca aggatcttac cgctgttgag atccagttcg atgtaacca ctcgtgcacc	2400
caactgatct tcagcatctt ttactttcac cagcgtttct gggtgagcaa aaacaggaag	2460
gcaaaatgcc gcaaaaaagg gaataagggc gacacggaaa tggtgaatac tcatactctt	2520
cctttttcaa tattattgaa gcatttatca gggttattgt ctcagagcg gatacatatt	2580
tgaatgtatt tagaaaaata aacaaatagg ggttccgcgc acatttcccc gaaaagtgcc	2640
acctgacgcg ccctgtagcg gcgcattaag cgcggcgggt gtggtggtta cgcgagcgt	2700
gaccgctaca cttgccagcg ccctagcgcc cgctcctttc gctttcttcc cttcctttct	2760
cgccacgttc gccggctttc cccgtcaagc tctaaatcgg gggctccctt taggggtccg	2820
atttagtgct ttacggcacc tcgaccccaa aaaacttgat tagggtgatg gttcacgtag	2880
tgggcatcg ccctgataga cggtttttcg ccctttgacg ttggagtcca cgttctttta	2940

025CIP4 PCT SEQ List.txt

tagtggactc	ttgtttccaaa	ctggaacaac	actcaaccct	atctcgggtct	attctttttga	3000
tttataaggg	atthttgccga	tttcggccta	ttggttaaaa	aatgagctga	tttaacaaaa	3060
atttaacgcg	aatttttaaca	aaatattaac	gtttacaatt	tcccattcgc	cattcaggct	3120
gcgcaactgt	tgggaagggc	gatcggtgcg	ggcctcttcg	ctattacgcc	agcccaagct	3180
accatgataa	gtaagtaata	ttaaggtagc	ggaggtagctt	ggagcggccg	caataaaata	3240
tctttattht	cattacatct	gtgtgtttgt	tttttgtgtg	aatcgatagt	actaacatac	3300
gctctccatc	aaaacaaaac	gaaacaaaac	aaactagcaa	aataggctgt	ccccagtgca	3360
agtgcagggtg	ccagaacatt	tctctatcga	taggtaccga	gctcttacgc	gtgctagccc	3420
tcgagcagga	tctatacatt	gaatcaatat	tggcaattag	ccatattagt	cattgggttat	3480
atagcataaa	tcaatattgg	ctattggcca	ttgcatacgt	tgtatctata	tcataaatatg	3540
tacattttata	ttggctcatg	tccaatatga	ccgccatgtt	gacattgatt	attgactagt	3600
tattaatagt	aatcaattac	gggggtcatta	gttcatagcc	catatatgga	gttccgcgtt	3660
acataactta	cggtaaatgg	cccgcctggc	tgaccgcca	acgacccccg	cccattgacg	3720
tcaataatga	cgatgtttcc	catagtaacg	ccaatagggg	ctttccattg	acgtcaatgg	3780
gtggaggtatt	tacggtaaac	tgcccacttg	gcagtacatc	aagtgtatca	tatgccaaagt	3840
ccgcccccta	ttgacgtcaa	tgacggtaaa	tggcccgcc	ggcattatgc	ccagtacatg	3900
accttacggg	actttcctac	ttggcagtag	atctacgtat	tagtcatcgc	tattaccatg	3960
gtgatgcgggt	tttggcagta	catcaatggg	cgtggatagc	ggtttgactc	acggggattt	4020
ccaagtctcc	accccatthga	cgtcaatggg	agtttgtttt	ggcaccaaaa	tcaacggggac	4080
tttccaaaat	gtcgtataca	ctccgcccc	ttgacgcaaa	tgggcggttag	gcgtgtacgg	4140
tgggaggtct	atataagcag	agctcgttta	gtgaaccgtc	agatcgccctg	gagacgccat	4200
ccacgctgtt	ttgacctcca	tagaagacac	cgggaccgat	ccagcctccc	ctcgaagctc	4260
gactctaggg	gctcgagatc	tgcatctaa	gtaagcttgc	atgcctgcag	gtcggccgcc	4320
acgaccgggtg	ccgccaccat	cccctgacct	acgcccctga	cccctcaca	ggagacgacc	4380
ttccatgacc	gagtacaagc	ccacggtgcg	cctcgccacc	cgcgacgacg	tccccggggc	4440
cgtacgcacc	ctcgccgccg	cgttcgccga	ctaccccgcc	acgcgccaca	ccgtcgacct	4500
ggaccgccac	atcgagcggg	tcaccgagct	gcaagaactc	ttcctcacgc	gcgtcgggct	4560
cgacatcggc	aagggtgtggg	tcgaggacga	cggcgccgcg	gtggcggtct	ggaccacgcc	4620
ggagagcgtc	gaagcggggg	cgggtgttcgc	cgagatcggc	ccgcgcatgg	ccgagttgag	4680
cggttccccg	ctggccgcgc	agcaacagat	ggaaggcctc	ctggcgccgc	accggcccaa	4740
ggagcccgcg	tggttcctgg	ccaccgtcgg	cgtctcgccc	gaccaccagg	gcaagggtct	4800
gggcagcgcc	gtcgtgctcc	ccggagtggg	ggcgcccgag	cgcgccgggg	tggccgcctt	4860

025CIP4 PCT SEQ List.txt

cctggagacc tccgcgcccc gcaacctccc cttctacgag cggctcggct tcaccgtcac	4920
cgccgacgtc gaggtgcccc aaggaccgcg cacctggtgc atgaccgcga agcccgggtgc	4980
ctgacgcccc cccacgacc cgacgcgccc gaccgaaagg agcgcacgac cccatggctc	5040
cgaccgaagc cgacccgggc ggccccgccc accccgcacc cgcccccgag gccaccgac	5100
t	5101

<210> 5
 <211> 5043
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> Plasmid pCMV-pur-attP

<400> 5	
ctagagtcgg ggcgccggc cgcttcgagc agacatgata agatacattg atgagtttgg	60
acaaaccaca actagaatgc agtgaaaaaa atgctttatt tgtgaaattt gtgatgctat	120
tgctttattt gtaaccatta taagctgcaa taaacaagtt aacaacaaca attgcattca	180
ttttatgttt caggttcagg gggagggtgtg ggaggttttt taaagcaagt aaaacctcta	240
caaatgtggt aaaatcgata aggatcaatt cggcttcgac tagtactgac ggacacaccg	300
aagccccggc ggcaaccctc agcggatgcc ccggggcttc acgttttccc aggtcagaag	360
cggttttcgg gagtagtgcc ccaactgggg taacctttga gttctctcag ttgggggcgt	420
agggtcgccg acatgacaca aggggttggtg accgggggtg acacgtacgc ggggtgcttac	480
gaccgtcagt cgcgcgagcg cgactagtag aagccgaatt gatccgtcga ccgatgccct	540
tgagagcctt caaccagtc agtccttcc ggtgggcgcg gggcatgact atcgtcgccg	600
cacttatgac tgtcttctt atcatgcaac tcgtaggaca ggtgccggca gcgctcttcc	660
gcttcctcgc tactgactc gctgcgctcg gtcgttcggc tgcggcgagc ggtatcagct	720
cactcaaagg cggaatacag gttatccaca gaatcagggg ataacgcagg aaagaacatg	780
tgagcaaaag gccagcaaaa ggccaggaac cgtaaaaagg ccgcgttgct ggcgtttttc	840
cataggctcc gccccctga cgagcatcac aaaaatcgac gctcaagtca gaggtggcga	900
aaccgcacag gactataaag ataccaggcg tttccccctg gaagctccct cgtgcgctct	960
cctgttccga ccctgccgct taccggatac ctgtccgcct ttctcccttc ggggaagcgtg	1020
gcgctttctc aatgctcacg ctgtaggtat ctgagttcgg ttaggtcgt tcgctccaag	1080
ctgggctgtg tgcacgaacc cccggttcag cccgaccgct gcgccttatc cggttaactat	1140
cgtcttgagt ccaaccgggt aagacacgac ttatcgccac tggcagcagc cactggtaac	1200
aggattagca gagcgaggta tgtaggcggt gctacagagt tcttgaagtg gtggcctaac	1260

025CIP4 PCT SEQ List.txt

tacggctaca	ctagaaggac	agtatttggg	atctgcgctc	tgctgaagcc	agttaccttc	1320
ggaaaaagag	ttggtagctc	ttgatccggc	aaacaaacca	ccgctggtag	cgggtggtttt	1380
tttgtttgca	agcagcagat	tacgcgcaga	aaaaaaggat	ctcaagaaga	tcctttgatc	1440
ttttctacgg	ggtctgacgc	tcagtggaac	gaaaactcac	gttaagggat	tttggtcatg	1500
agattatcaa	aaaggatcct	cacctagatc	cttttaaatt	aaaaatgaag	ttttaaatca	1560
atctaaagta	tatatgagta	aacttgggtc	gacagttacc	aatgcttaat	cagtgaggca	1620
cctatctcag	cgatctgtct	atttcgttca	tccatagttg	cctgactccc	cgctcgttag	1680
ataactacga	tacgggaggg	cttaccatct	ggccccagtg	ctgcaatgat	accgcgagac	1740
ccacgctcac	cggctccaga	tttatcagca	ataaaccagc	cagccggaag	ggccgagcgc	1800
agaagtggtc	ctgcaacttt	atccgcctcc	atccagtcta	ttaattgttg	ccgggaagct	1860
agagtaagta	gttcgccagt	taatagtttg	cgcaacgttg	ttgccattgc	tacaggcatc	1920
gtggtgtcac	gctcgtcgtt	tggtatggct	tcattcagct	ccggttccca	acgatcaagg	1980
cgagttacat	gatcccccat	gttgtgcaaa	aaagcggtta	gtccttcggg	tcctccgatc	2040
gttgtcagaa	gtaagttggc	cgcagtgtta	tcactcatgg	ttatggcagc	actgcataat	2100
tctcttactg	tcatgccatc	cgtaagatgc	ttttctgtga	ctggtgagta	ctcaaccaag	2160
tcattctgag	aatagtgtat	gcggcgaccg	agttgctctt	gcccggcgctc	aatacgggat	2220
aataccgcgc	cacatagcag	aacttttaaa	gtgctcatca	ttggaaaacg	ttcttcgggg	2280
cgaaaactct	caaggatcct	accgctgttg	agatccagtt	cgatgtaacc	cactcgtgca	2340
cccaactgat	cttcagcatc	ttttactttc	accagcgttt	ctgggtgagc	aaaaacagga	2400
aggcaaaatg	ccgcaaaaaa	gggaataagg	gcgacacgga	aatgttgaat	actcatactc	2460
ttcctttttc	aatattattg	aagcatttat	cagggttatt	gtctcatgag	cggatacata	2520
tttgaatgta	tttagaaaaa	taaacaaata	ggggttccgc	gcacatttcc	ccgaaaagtg	2580
ccacctgacg	cgccctgtag	cggcgcatca	agcgcgggcg	gtgtgggtgg	tacgcgcagc	2640
gtgaccgcta	cacttgccag	cgccctagcg	cccgtcctt	tcgctttctt	cccttccttt	2700
ctcgccacgt	tcgccgggctt	tccccgtcaa	gctctaaatc	gggggctccc	tttagggttc	2760
cgatttagtg	ctttacggca	cctcgacccc	aaaaaacttg	attagggatga	tggttcacgt	2820
agtgggcat	cgccctgata	gacggttttt	cgccctttga	cgttggagtc	cacgttcttt	2880
aatagtggac	tcttgttcca	aactggaaca	acactcaacc	ctatctcggt	ctattctttt	2940
gatttataag	ggatttttgc	gatttcggcc	tatttggttaa	aaaatgagct	gatttaacaa	3000
aaatttaacg	cgaattttta	caaaatatta	acgtttacaa	tttcccattc	gccattcagg	3060
ctgcgcaact	gttgggaagg	gcgatcggtg	cgggcctctt	cgctattacg	ccagcccaag	3120
ctaccatgat	aagtaagtaa	tattaaggta	cgggaggtac	ttggagcggc	cgcaataaaa	3180

025CIP4 PCT SEQ List.txt

tatcttttatt	ttcattacat	ctgtgtgttg	gttttttgtg	tgaatcgata	gtactaacat	3240
acgctctcca	tcaaaacaaa	acgaaacaaa	acaaactagc	aaaataggct	gtccccagt	3300
caagtgcagg	tgccagaaca	tttctctatc	gataggtagc	gagctcttac	gcgtgctagc	3360
cctcgagcag	gatctataca	ttgaatcaat	attggcaatt	agccatatta	gtcattgggt	3420
atatagcata	aatcaatatt	ggctattggc	cattgcatac	gttgatatcta	tatcataata	3480
tgtacattta	tattggctca	tgtccaatat	gaccgccatg	ttgacattga	ttattgacta	3540
gttattaata	gtaatcaatt	acgggggtcat	tagttcatag	cccatatatg	gagttccgcg	3600
ttacataact	tacggtaa	ggccccgcctg	gctgaccgcc	caacgacccc	cgccccattga	3660
cgtcaataat	gacgtatgtt	cccatagtaa	cgccaatagg	gactttccat	tgacgtcaat	3720
gggtggagta	tttacggtaa	actgcccact	tggcagtaca	tcaagtgtat	catatgccaa	3780
gtccgcccc	tattgacgtc	aatgacggta	aatggccccg	ctggcattat	gcccagtaca	3840
tgaccttacg	ggactttcct	acttggcagt	acatctacgt	attagtcac	gctattacca	3900
tggtgatg	gttttggcag	tacatcaatg	ggcgtggata	gcggtttgac	tcacggggat	3960
ttccaagtct	ccacccatt	gacgtcaatg	ggagtttgtt	ttggcaccaa	aatcaacggg	4020
actttccaaa	atgtcgtaac	aactccgccc	cattgacgca	aatgggcggg	aggcgtgtac	4080
ggtgggaggt	ctatataagc	agagctcggt	tagtgaaccg	tcagatcgcc	tggagacgcc	4140
atccacgctg	ttttgacctc	catagaagac	accgggaccg	atccagcctc	ccctcgaagc	4200
tcgactctag	gggctcgaga	tctgcgatct	aagtaagctt	gcatgcctgc	aggctggccg	4260
ccacgaccgg	tgccgccacc	atccccctgac	ccacgcccct	gacccctcac	aaggagacga	4320
ccttccatga	ccgagtacaa	gcccacgggtg	cgccctcgca	cccgcgacga	cgtcccccg	4380
gccgtacgca	ccctcgccgc	cgcgttcgcc	gactaccccg	ccacgcgcca	caccgtcgac	4440
ccggaccgcc	acatcgagcg	ggtcaccgag	ctgcaagaac	tcttctctac	gcgcgtcggg	4500
ctcgacatcg	gcaaggtgtg	ggtcgcggac	gacggcgccg	cggtggcggt	ctggaccacg	4560
ccggagagcg	tcgaagcggg	ggcgggtgtt	gccgagatcg	gcccgcgcat	ggccgagttg	4620
agcgggtccc	ggctggccgc	gcagcaacag	atggaaggcc	tcctggcgcc	gcaccggccc	4680
aaggagcccc	cgtggttctt	ggccaccgtc	ggcgtctcgc	ccgaccacca	gggcaagggt	4740
ctgggcagcg	ccgtcgtgct	ccccggagtg	gaggcgcccg	agcgcgccgg	ggtgcccgcc	4800
ttcctggaga	cctccgcgcc	ccgcaacctc	cccttctacg	agcggctcgg	cttcaccgtc	4860
accgccgacg	tcgaggtgcc	cgaaggaccg	cgcacctggg	gcatgaccgg	caagccccgt	4920
gcctgacgcc	cgccccacga	cccgacgcgc	ccgaccgaaa	ggagcgcacg	accccatggc	4980
tccgaccgaa	gccgaccggg	gcggccccgc	cgaccccgca	cccgcccccg	aggcccaccg	5040

025CIP4 PCT SEQ List.txt

act

5043

<210> 6
 <211> 5041
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> Plasmid pCMV-EGFP-attB

<400> 6
 ctagagtcg ggcggccggc cgcttcgagc agacatgata agatacattg atgagtttgg 60
 acaaaccaca actagaatgc agtgaaaaaa atgctttatt tgtgaaattt gtgatgctat 120
 tgctttatatt gtaaccatta taagctgcaa taaacaagtt aacaacaaca attgcattca 180
 ttttatgttt cagggttcagg gggaggtgtg ggaggttttt taaagcaagt aaaacctcta 240
 caaatgtggt aaaatcgata aggatcaatt cggcttcagg taccgtcgac gatgtaggtc 300
 acggtctcga agccgcggtg cgggtgccag ggcgtgccct tgggctcccc gggcgcgtac 360
 tccacctcac ccatctgggtc catcatgatg aacgggtcga ggtggcggta gttgatcccg 420
 gcgaacgcgc ggcgcaccgg gaagccctcg ccttcgaaac cgctgggcgc ggtgggtcacg 480
 gtgagcacgg gacgtgcgac ggcgtcggcg ggtgcggata cgcggggcag cgtcagcggg 540
 ttctcgacgg tcacggcggg catgtcgaca gccgaattga tccgtcgacc gatgcccttg 600
 agagccttca acccagtcag ctccctccgg tgggcgcggg gcatgactat cgtcgccgca 660
 cttatgactg tcttctttat catgcaactc gtaggacagg tgccggcagc gctcttccgc 720
 ttctcgctc actgactcgc tgcgctcggg cgttcggctg cggcgagcgg tatkagctca 780
 ctcaaaggcg gtaatacggg tatccacaga atcaggggat aacgcaggaa agaacatgtg 840
 agcaaaaggc cagcaaaagg ccaggaaccg taaaaaggcc gcgttgctgg cgtttttcca 900
 taggctccgc ccccttgacg agcatcaca aaatcgacgc tcaagtcaga ggtggcgaaa 960
 cccgacagga ctataaagat accaggcggt tccccctgga agctccctcg tgcgctctcc 1020
 tgttccgacc ctgccgctta ccggatacct gtccgccttt ctcccttcgg gaagcgtggc 1080
 gctttctcaa tgctcacgct gtaggtatct cagttcgggtg taggtcgttc gctccaagct 1140
 gggctgtgtg cacgaacccc ccgttcagcc cgaccgtgc gccttatccg gtaactatcg 1200
 tcttgagtcc aaccgggtaa gacacgactt atcgccactg gcagcagcca ctggtaacag 1260
 gattagcaga gcgaggtatg taggcgggtg tacagagttc ttgaagtggg ggcctaacta 1320
 cggctacact agaaggacag tatttggtat ctgcgctctg ctgaagccag ttaccttcgg 1380
 aaaaagagtt ggtagctctt gatccggcaa acaaaccacc gctggtagcg gtgggtttttt 1440
 tgtttgcaag cagcagatta cgcgcagaaa aaaaggatct caagaagatc ctttgatctt 1500
 ttctacgggg tctgacgctc agtggaacga aaactcacgt taagggattt tgggtcatgag 1560

025CIP4 PCT SEQ List.txt

attatcaaaa	aggatcttca	cctagatcct	tttaaattaa	aaatgaagtt	ttaaatcaat	1620
ctaaagtata	tatgagtaaa	cttgggtctga	cagttaccaa	tgcttaataca	gtgaggcacc	1680
tatctcagcg	atctgtctat	ttcgtttcatc	catagttgcc	tgactccccg	tcgtgtagat	1740
aactacgata	cgggagggct	taccatctgg	ccccagtgt	gcaatgatac	cgcgagaccc	1800
acgctcaccg	gctccagatt	tatcagcaat	aaaccagcca	gccggaaggg	ccgagcgcag	1860
aagtgggtcct	gcaactttat	ccgcctccat	ccagtctatt	aattgttgcc	gggaagctag	1920
agtaagtagt	tcgccagtta	atagtttgcg	caacgttggt	gccattgcta	caggcatcgt	1980
ggtgtcacgc	tcgtcgtttg	gtatggcttc	attcagctcc	ggttcccaac	gatcaaggcg	2040
agttacatga	tcctccatgt	tgtgcaaaaa	agcggtttagc	tccttcggtc	ctccgatcgt	2100
tgtcagaagt	aagttggccg	cagtgttatc	actcatgggt	atggcagcac	tgcataattc	2160
tcttactgtc	atgccatccg	taagatgctt	ttctgtgact	ggtgagtact	caaccaagtc	2220
attctgagaa	tagtgatatgc	ggcgaccgag	ttgctcttgc	ccggcgtcaa	tacgggataa	2280
taccgcgcca	catagcagaa	ctttaaaagt	gctcatcatt	ggaaaacggt	cttcggggcg	2340
aaaactctca	aggatcttac	cgctgttgag	atccagttcg	atgtaaccca	ctcgtgcacc	2400
caactgatct	tcagcatctt	ttactttcac	cagcgtttct	gggtgagcaa	aaacaggaag	2460
gcaaaatgcc	gcaaaaaagg	gaataagggc	gacacggaaa	tggtgaatac	tcataactctt	2520
cctttttcaa	tattattgaa	gcatttatca	gggttattgt	ctcatgagcg	gatacatatt	2580
tgaatgtatt	tagaaaaata	aacaaatagg	ggttccgcgc	acatttcccc	gaaaagtgcc	2640
acctgacgcg	ccctgtagcg	gcgcatthaag	cgcgggcggt	gtggtgggta	cgcgcagcgt	2700
gaccgctaca	cttgccagcg	ccctagcgcc	cgctcctttc	gctttcttcc	cttcctttct	2760
cgccacgttc	gccggctttc	cccgtcaagc	tctaaatcgg	gggctccctt	taggggttccg	2820
atttagtgct	ttacggcacc	tcgaccccaa	aaaacttgat	taggggtgatg	gttcacgtag	2880
tgggccatcg	ccctgataga	cggtttttcg	ccctttgacg	ttggagtcca	cgttctttaa	2940
tagtggactc	ttgttccaaa	ctggaacaac	actcaaccct	atctcgggtct	attcttttga	3000
tttataaggg	atthttgccga	tttcggccta	ttgggttaaaa	aatgagctga	tttaacaaaa	3060
atttaacgcg	aattttaaca	aaatattaac	gtttacaatt	tcccatcgc	cattcaggct	3120
gcgcaactgt	tgggaagggc	gatcgggtgcg	ggcctcttcg	ctattacgcc	agcccaagct	3180
accatgataa	gtaagtaata	ttaaggtacg	ggaggtactt	ggagcggccg	caataaaata	3240
tctttattht	cattacatct	gtgtgttggt	tttttgtgtg	aatcgatagt	actaacatac	3300
gctctccatc	aaaacaaaac	gaaacaaaac	aaactagcaa	aataggctgt	ccccagtgc	3360
agtgacgggtg	ccagaacatt	tctctatcga	taggtaccga	gctcttacgc	gtgctagccc	3420

025CIP4 PCT SEQ List.txt

tcgagcagga tctatacatt gaatcaatat tggcaattag ccatattagt cattggttat	3480
atagcataaa tcaatatttg ctattggcca ttgcatacgt tgtatctata tcataaatatg	3540
tacattttata ttggctcatg tccaatatga ccgccatgtt gacattgatt attgactagt	3600
tattaatagt aatcaattac ggggtcatta gttcatagcc catatatgga gttccgcgtt	3660
acataactta cggtaaattg cccgcctggc tgaccgcca acgacccccg cccattgacg	3720
tcaataatga cgtatgttcc catagtaacg ccaataggga ctttccattg acgtcaatgg	3780
gtggagtatt tacggtaaac tgcccacttg gcagtacatc aagtgtatca tatgccaaagt	3840
ccgcccccta ttgacgtcaa tgacggtaaa tggcccgctt ggcattatgc ccagtacatg	3900
accttacggg acttttctac ttggcagtac atctacgtat tagtcatcgc tattaccatg	3960
gtgatgcggt tttggcagta catcaatggg cgtggatagc ggtttgactc acggggattt	4020
ccaagtctcc accccattga cgtcaatggg agtttgtttt ggcacaaaaa tcaacgggac	4080
tttccaaaat gtcgtaacaa ctccgcccc a ttgacgcaa tgggcggtag gcgtgtacgg	4140
tgggaggtct atataagcag agctcgttta gtgaaccgtc agatcgccctg gagacgcat	4200
ccacgtgtt ttgacctca tagaagacac cgggaccgat ccagcctccc ctggaagctc	4260
gactctaggg gctcgagatc cccgggtacc ggtcgccacc atggtgagca agggcgagga	4320
gctgttcacc ggggtggtgc ccatcctggt cgagctggac ggcgacgtaa acggccacaa	4380
gttcagcgtg tccggcgagg gcgagggcga tgccacctac ggcaagctga ccctgaagtt	4440
catctgcacc accggcaagc tgcccgtgcc ctggcccacc ctctgtacca ccctgacct	4500
cggcgtgcag tgcttcagcc gctaccccga ccacatgaag cagcacgact tcttcaagtc	4560
cgccatgccc gaaggctacg tccaggagcg caccatcttc ttcaaggacg acggcaacta	4620
caagaccgc gccgaggtga agttcgagg cgacaccctg gtgaaccgca tcgagctgaa	4680
gggcatcgac ttcaaggagg acggcaacat cctggggcac aagctggagt acaactacaa	4740
cagccacaac gtctatatca tggccgacaa gcagaagaac ggcataagg tgaacttcaa	4800
gatccgccac aacatcgagg acggcagcgt gcagctcgcc gaccactacc agcagaacac	4860
ccccatcggc gacggccccg tgctgctgcc cgacaaccac tacctgagca cccagtccgc	4920
cctgagcaaa gacccaacg agaagcgca tcacatggtc ctgctggagt tcgtgaccgc	4980
cgccgggatc actctcggca tggacgagct gtacaagtaa agcggccgct cgagcatgca	5040
t	5041

<210> 7

<211> 18116

<212> DNA

<213> Artificial sequence

<220>

025CIP4 PCT SEQ List.txt

<223> Plasmid p12.0_{lys}-LSPIPMM-CMV-pur-attB

<400> 7

gggctgcagg aattcgattg ccgccttctt tgatattcac tctgttgat ttcattctctt	60
cttgccgatg aaaggatata acagtctgta taacagtctg tgaggaaata cttggtatctt	120
cttctgatca gtgtttttat aagtaatgtt gaatattgga taaggctgtg tgtcctttgt	180
cttgggagac aaagcccaca gcagggtggtg gttgggggtg tggcagctca gtgacaggag	240
aggttttttt gcctgttttt tttttttttt ttttttttaa gtaagggtgtt cttttttctt	300
agtaaatctt ctactggact gtatgttttg acaggtcaga aacatttctt caaaagaaga	360
accttttgga aactgtacag cccttttctt tcattccctt tttgctttct gtgccaatgc	420
ctttggttct gattgcatta tggaaaacgt tgatcggaac ttgaggtttt tatttatagt	480
gtggcttgaa agcttgata gctgttgta cagagatac cttattaagt ttaggccagc	540
ttgatgcttt attttttccc tttgaagtag tgagcgttct ctggtttttt tcctttgaaa	600
ctggtgaggc ttagatcttt ctaatgggat tttttacctg atgatctagt tgcataccca	660
aatgcttgta aatgttttcc tagttaacat gttgataact tcggattttac atgttgata	720
tacttgcat ctgtgtttct agtaaaaata tatggcattt atagaaatac gtaattcctg	780
atttcctttt tttttatctc tatgctctgt gtgtacaggt caaacagact tcactcctat	840
ttttatttat agaattttat atgcagtctg tcgttggttc ttgtgttgta aggatacagc	900
cttaaatttc ctagagcgat gctcagtaag gcgggttgtc acatgggttc aaatgtaaaa	960
cgggcacgtt tggctgctgc cttcccagaga tccaggacac taaactgctt ctgcactgag	1020
gtataaatcg cttcagatcc cagggaagtg cagatccacg tgcataattct taaagaagaa	1080
tgaatacttt ctaaaatatt ttggcatagg aagcaagctg catggatttg tttgggactt	1140
aaattatttt ggtaacggag tgcataagggt tttaaacacag ttgcagcatg ctaacgagtc	1200
acagcgttta tgcagaagtg atgcctggat gcctgttgca gctgtttacg gcactgcctt	1260
gcagttagca ttgcagatag ggggtgggtg ctttgtgtcg tgttcccaca cgctgccaca	1320
cagccacctc ccggaacaca tctcacctgc tgggtacttt tcaaaccatc ttagcagtag	1380
tagatgagtt actatgaaac agagaagttc ctcagttgga tattctcatg ggatgtcttt	1440
tttcccatgt tgggcaaagt atgataaagc atctctatct gtaaattatg cacttgtag	1500
ttcctgaatc ctttctatag caccacttat tgcagcaggt gtaggctctg gtgtggcctg	1560
tgtctgtgct tcaatctttt aaagcttctt tggaaataca ctgacttgat tgaagtctct	1620
tgaagatagt aaacagtact tacctttgat cccaatgaaa tcgagcattt cagttgtaaa	1680
agaattccgc ctattcatac catgtaatgt aattttacac cccagtgct gacactttgg	1740
aatatattca agtaatagac tttggcctca ccctcttggt tactgtatct tgtaatagaa	1800

025CIP4 PCT SEQ List.txt

aatatttttaa	actgtgcata	tgattattac	attatgaaag	agacattctg	ctgatcttca	1860
aatgtaagaa	aatgaggagt	gcgtgtgctt	ttataaatac	aagtgattgc	aaattagtg	1920
aggtgtcctt	aaaaaaaaa	aaaaaaagta	atataaaaag	gaccaggtgt	tttacaagtg	1980
aaatacattc	ctatttggt	aacagttaca	tttttatgaa	gattaccagc	gctgctgact	2040
ttctaaacat	aaggctgtat	tgtcttcctg	taccattgca	tttcctcatt	cccaatttgc	2100
acaaggatgt	ctgggtaaac	tattcaagaa	atggctttga	aatacagcat	gggagcttgt	2160
ctgagttgga	atgcagagtt	gcactgcaaa	atgtcaggaa	atggatgtct	ctcagaatgc	2220
ccaactccaa	aggattttat	atgtgtatat	agtaagcagt	ttcctgattc	cagcaggcca	2280
aagagtctgc	tgaatgttgt	gttgccggag	acctgtat	ctcaacaagg	taagatggta	2340
tcctagcaac	tgcggtttt	aatacatttt	cagcagaagt	acttagtta	tctctacctt	2400
tagggatcgt	ttcatcattt	ttagatgtta	tacttgaaat	actgcataac	ttttagcttt	2460
catgggttcc	tttttttcag	cctttaggag	actgttaagc	aatttgctgt	ccaacttttg	2520
tgttggctct	aaactgcaat	agtagtttac	cttgtattga	agaaataaag	accattttta	2580
tattaaaaaa	tacttttgtc	tgtcttcatt	ttgacttgct	tgatatcctt	gcagtgccca	2640
ttatgtcagt	tctgtcagat	attcagacat	caaaacttaa	cgtgagctca	gtggagttac	2700
agctgcgggt	ttgatgctgt	tattatttct	gaaactagaa	atgatgttgt	cttcatctgc	2760
tcatcaaaca	cttcatgcag	agtgtgaagg	tagtgagaaa	tgatacatt	tattgatact	2820
tttttaaaagt	caacttttta	tcagattttt	ttttcatttg	gaaatatatt	gttttctaga	2880
ctgcatagct	tctgaatctg	aaatgcagtc	tgattggcat	gaagaagcac	agcactcttc	2940
atcttactta	aacttcattt	tggaatgaag	gaagttaagc	aagggcacag	gtccatgaaa	3000
tagagacagt	gcgctcagga	gaaagtgaac	ctggatttct	ttggctagt	ttctaaatct	3060
gtagtgagga	aagtaacacc	cgattccttg	aaagggctcc	agctttaatg	cttccaaatt	3120
gaaggtggca	ggcaacttgg	ccactgggtta	tttactgcat	tatgtctcag	tttcgcagct	3180
aacctggctt	ctccactatt	gagcatggac	tatagcctgg	cttcagaggc	caggtgaagg	3240
ttgggatggg	tggaaggagt	gctgggctgt	ggctgggggg	actgtgggga	ctccaagctg	3300
agcttggggg	gggcagcaca	gggaaaagt	tgggtaacta	tttttaagta	ctgtgttgca	3360
aacgtctcat	ctgcaaatac	gtaggggtgt	tactctcgaa	gattaacagt	gtgggttcag	3420
taatatatgg	atgaattcac	agtgggaagca	ttcaagggtta	gatcatctaa	cgacaccaga	3480
tcatcaagct	atgattggaa	gcggtatcag	aagagcgagg	aaggtaagca	gtcttcatat	3540
gttttcctc	cacgtaaagc	agtctgggaa	agtagcacc	cttgagcaga	gacaaggaaa	3600
taattcagga	gcatgtgcta	ggagaacttt	cttgctgaat	tctacttgca	agagctttga	3660
tgcctggctt	ctggtgcctt	ctgcagcacc	tgcaaggccc	agagcctgtg	gtgagctgga	3720

O25CIP4 PCT SEQ List.txt

gggaaagatt	ctgctcaagt	ccaagcttca	gcaggtcatt	gtctttgctt	cttccccag	3780
cactgtgcag	cagagtggaa	ctgatgtcga	agcctcctgt	ccactacctg	ttgctgcagg	3840
cagactgctc	tcagaaaaag	agagctaact	ctatgccata	gtctgaagggt	aaaatggggt	3900
ttaaaaaaga	aaacacaaag	gcaaaaccgg	ctgccccatg	agaagaaagc	agtggtaaac	3960
atggtagaaa	aggtgcagaa	gcccccaggc	agtgtgacag	gcccctcctg	ccacctagag	4020
gcgggaacaa	gcttccctgc	ctagggctct	gcccgcgaag	tgctgttttc	tttgggtgggt	4080
tttgtttggc	gtttggtttt	gagatttaga	cacaagggaa	gcctgaaagg	aggtgttggg	4140
cactattttg	gtttgtaaag	cctgtacttc	aaatatatat	tttgtgaggg	agtgtagcga	4200
attggccaat	ttaaaataaa	gttgcaagag	attgaaggct	gagtagttga	gagggttaaca	4260
cgtttaatga	gatcttctga	aactactgct	tctaaacact	tgtttgagtg	gtgagacctt	4320
ggataggtga	gtgctcttgt	tacatgtctg	atgcacttgc	ttgtcctttt	ccatccacat	4380
ccatgcattc	cacatccacg	catttgtcac	ttatcccata	tctgtcatat	ctgacatacc	4440
tgtctcttcg	tcacttggtc	agaagaaaca	gatgtgataa	tccccagccg	ccccaagtgt	4500
gagaagatgg	cagttgcttc	tttccctttt	tcctgctaag	taaggatttt	ctcctggctt	4560
tgacacctca	cgaaatagtc	ttcctgcctt	acattctggg	cattatttca	aatatctttg	4620
gagtgcgctg	ctctcaagtt	tgtgtcttcc	tactcttaga	gtgaatgctc	ttagagtga	4680
agagaaggaa	gagaagatgt	tggccgcagt	tctctgatga	acacacctct	gaataatggc	4740
caaagggtggg	tgggtttctc	tgaggaacgg	gcagcgtttg	cctctgaaag	caaggagctc	4800
tgcgagattg	cagttatttt	gcaactgatg	gtggaactgg	tgcttaaagc	agattcccta	4860
ggttccctgc	tacttctttt	ccttcttggc	agtcagttta	tttctgacag	acaaacagcc	4920
acccccactg	caggcttaga	aagtatgtgg	ctctgcctgg	gtgtgttaca	gctctgccct	4980
ggtgaaaggg	gattaaaacg	ggcaccattc	atcccaaaca	ggatcctcat	tcatggatca	5040
agctgtaagg	aacttgggct	ccaacctcaa	aacattaatt	ggagtacgaa	tgtaattaaa	5100
actgcattct	cgcattccta	agtcatttag	tctggactct	gcagcatgta	ggtcggcagc	5160
tcccactttc	tcaaagacca	ctgatggagg	agtagtaaaa	atggagaccg	attcagaaca	5220
accaacggag	tgttgccgaa	gaaactgatg	gaaataatgc	atgaattgtg	tggtggacat	5280
ttttttttaa	tacataaact	acttcaaata	aggtcggaga	aggtcagtg	tttattagca	5340
gccataaaac	cagggtgagcg	agtaccattt	ttctctacaa	gaaaaacgat	tctgagctct	5400
gcgtaagtat	aagttctcca	tagcggctga	agctcccccc	tggctgcctg	ccatctcagc	5460
tggagtgcag	tgccatttcc	ttgggggttt	tctcacagca	gtaatgggac	aatacttcac	5520
aaaaattctt	tcttttcctg	tcatgtggga	tccctactgt	gccctcctgg	ttttacgtta	5580

025CIP4 PCT SEQ List.txt

ccccctgact gttccattca gcggttttggg aagagaaaaa gaatttggaa ataaaacatg	5640
tctacgttat cacctcctcc agcatttttg tttttaatta tgtcaataac tggcttagat	5700
ttggaaatga gagggggttg ggtgtattac cgaggaacaa aggaaggctt atataaactc	5760
aagtctttta tttagagaac tggcaagctg tcaaaaacaa aaaggcctta ccaccaaatt	5820
aagtgaatag ccgctatagc cagcagggcc agcacgaggg atggtgcact gctggcacta	5880
tgccacggcc tgcttgtgac tctgagagca actgcttttg aaatgacagc acttgggtgca	5940
atttcctttg tttcagaatg cgtagagcgt gtgcttggcg acagtttttc tagttaggcc	6000
acttcttttt tccttctctc ctcatctctc taagcatgtc tccatgctgg taatcccagt	6060
caagtgaacg ttcaaacaat gaatccatca ctgtaggatt ctctgtggtga tcaaactttt	6120
gtgtgaggtc tataaaatat ggaagcttat ttatttttcg ttcttccata tcagtcttct	6180
ctatgacaat tcacatccac cacagcaaat taaagggtgaa ggaggctggg gggatgaaga	6240
gggtcttcta gctttacgtt cttccttgca aggccacagg aaaatgctga gagctgtaga	6300
atacagcctg gggtaagaag ttcagtctcc tgctgggaca gctaaccgca tcttataacc	6360
ccttctgaga ctcatcttag gaccaaatag ggtctatctg gggtttttgt tcctgctgtt	6420
cctcctggaa ggctatctca ctatttcact gctcccacgg ttacaaacca aagatacagc	6480
ctgaattttt tctaggccac attacataaa ttgacctgg taccaatatt gttctctata	6540
tagttatttc cttccccact gtgtttaacc ccttaaggca ttcagaacaa ctagaatcat	6600
agaatggttt ggattggaag gggccttaaa catcatccat ttccaaccct ctgccatggg	6660
ctgcttgcca cccactggct caggctgccc agggcccat ccagcctggc cttgagcacc	6720
tccagggatg gggcaccac agcttctctg ggcagcctgt gccaacacct caccactctc	6780
tgggtaaaga attctctttt aacatctaata ctaaatctct tctcttttag tttaaagcca	6840
ttcctctttt tcccgttgct atctgtccaa gaaatgtgta ttggtctccc tcctgcttat	6900
aagcaggaag tactggaagg ctgcagtga gtctccccac agccttctct tctccaggct	6960
gaacaagccc agctccttca gcctgtcttc gtaggagatc atcttagtgg ccctcctctg	7020
gaccatttcc aacagttcca cggctttctt gtggagcccc aggtctggat gcagtacttc	7080
agatggggcc ttacaaaggc agagcagatg gggacaatcg cttaccctc cctgctggct	7140
gcccctgttt tgatgcagcc cagggtactg ttggcctttc aggtctccag accccttgct	7200
gatttgtgtc aagcttttca tccaccagaa cccacgcttc ctggttaata cttctgacct	7260
cacttctgta agcttgtttc aggagacttc cattcttttag gacagactgt gttacaccta	7320
cctgccctat tcttgcatat atacatttca gttcatgttt cctgtaacag gacagaatat	7380
gtattcctct aacaaaaata catgcagaat tcctagtgcc atctcagtag ggttttcatg	7440
gcagtattag cacatagtca atttgctgca agtaccttcc aagctgcggc ctcccataaa	7500

025CIP4 PCT SEQ List.txt

tcctgtat	ttt	gggatcag	tt	acctttt	ggg	gtaagct	tttt	gtatctgcag	agaccctggg	7560
ggttctgatg	tgcttcagct	ctgctctgtt	ctgactgcac	cattttctag	atcaccagct	7620				
tgttcctgta	caacttcctt	gtcctccatc	ctttcccagc	ttgtatcttt	gacaaataca	7680				
ggcctat	tttt	tgtgtttgct	tcagcagcca	tttaattctt	cagtgtcatc	ttgttctgtt	7740			
gatgccactg	gaacaggatt	ttcagcagtc	ttgcaaagaa	catctagctg	aaaactttct	7800				
gccattcaat	attctttacca	gttcttcttg	tttgaggtga	gccataaatt	actagaactt	7860				
cgtcactgac	aagtttatgc	attttattac	ttctattatg	tacttacttt	gacataacac	7920				
agacacgcac	atattttgct	gggatttcca	cagtgtctct	gtgtccttca	catggtttta	7980				
ctgtcatact	tccgttataa	ccttggcaat	ctgcccagct	gcccatacaca	agaaaagaga	8040				
ttcctttttt	attacttctc	ttcagccaat	aaacaaaatg	tgagaagccc	aaacaagaac	8100				
ttgtggggca	ggctgccatc	aagggagaga	cagctgaagg	gttgtgtagc	tcaatagaat	8160				
taagaaataa	taaagctgtg	tcagacagtt	ttgcctgatt	tatacaggca	cgccccaagc	8220				
cagagaggct	gtctgccaaag	gccaccttgc	agtccttggg	ttgtaagata	agtcataaggt	8280				
aacttttctg	gtgaattgcg	tggagaatca	tgatggcag	tcttgctgtt	tactatggta	8340				
agatgctaaa	ataggagaca	gcaaagtaac	acttgctgct	gtaggtgctc	tgctatccag	8400				
acagcgatgg	cactcgcaca	ccaagatgag	ggatgctccc	agctgacgga	tgctggggca	8460				
gtaacagtgg	gtcccatgct	gcctgctcat	tagcatcacc	tcagccctca	ccagcccatc	8520				
agaaggatca	tccaagctg	aggaaagt	ctcatcttct	tcacatcatc	aaacctttgg	8580				
cctgactgat	gcctcccga	tgcttaa	atg	tggtcactga	catctttatt	tttctatgat	8640			
ttcaagtcag	aacctccga	tcaggaggga	acacatagtg	ggaatgtacc	ctcagctcca	8700				
aggccagatc	ttccttcaat	gatcatgcat	gctacttagg	aagggtgtgtg	tgtgtgaatg	8760				
tagaattgcc	tttgttat	tttcttctg	ctgtcaggaa	cattttgaat	accagagaaa	8820				
aagaaaagtg	ctcttcttgg	catgggagga	gttgtcacac	ttgcaaaata	aaggatgcag	8880				
tcccaa	atgt	tcataatctc	agggctctgaa	ggaggatcag	aaactgtgta	tacaatttca	8940			
ggcttctctg	aatgcagctt	ttgaaagctg	ttcctggccg	aggcagtact	agtcagaacc	9000				
ctcgga	aaaca	atg	tcttcaaggt	gcagcaggag	gaaacacctt	gcccatactg	9060			
aaagtgaata	accactgccg	ctgaagggaat	ccagctcctg	tttgagcagg	tgctgcacac	9120				
tcccacactg	aaacaacagt	tcatTTTTAT	aggacttcca	ggaaggatct	tcttcttaag	9180				
cttcttaatt	atggtacatc	tccagttggc	agatgactat	gactactgac	aggagaatga	9240				
ggaactagct	gggaatat	ctgtttgacc	accatggagt	caccattt	tttactggta	9300				
tttggaata	ataattctga	attgcaaagc	aggagttagc	gaagatcttc	atttcttcca	9360				

025CIP4 PCT SEQ List.txt

tggttggtgac agcacagttc tggctatgaa agtctgctta caaggaagag gataaaaatc	9420
ataggggataa taaatctaag tttgaagaca atgaggtttt agctgcattt gacatgaaga	9480
aattgagacc tctactggat agctatggta tttacgtgtc tttttgctta gttacttatt	9540
gacccagct gaggtcaagt atgaactcag gtctctcggg ctactggcat ggattgatta	9600
catacaactg taattttagc agtgatttag ggtttatgag tacttttgca gtaaatcata	9660
gggttagtaa tgtaaatctc agggaaaaaa aaaaaaagcc aaccctgaca gacatcccag	9720
ctcaggtgga aatcaaggat cacagctcag tgcggtccca gagaacacag ggactcttct	9780
cttaggacct ttatgtacag ggcctcaaga taactgatgt tagtcagaag actttccatt	9840
ctggccacag ttcagctgag gcaatcctgg aattttctct ccgctgcaca gttccagtca	9900
tcccagtttg tacagtctg gcactttttg ggtcaggccg tgatccaagg agcagaagtt	9960
ccagctatgg tcagggagtg cctgaccgtc ccaactcact gcactcaaac aaaggcgaaa	10020
ccacaagagt ggcttttggt gaaattgcag tgtggcccag aggggctgca ccagtactgg	10080
attgaccacg aggcaacatt aatcctcagc aagtgcatt tgcagccatt aaattgaact	10140
aactgatact acaatgcaat cagtatcaac aagtggtttg gcttggaaga tggagtctag	10200
gggctctaca ggagtagcta ctctctaatt gagttgcatt ttgaagcagg aactgtgaa	10260
aagctggcct cctaaagagg ctgctaaaca ttaggggtcaa tttccagtg cactttctga	10320
agtgtctgca gttcccatg caaagctgcc caaacatagc acttccaatt gaatacaatt	10380
atatgcaggc gtactgctt ttgccagcac tgtccttctc aaatgaactc aacaacaat	10440
ttcaaagtct agtagaaagt aacaagcttt gaatgtcatt aaaaagtata tctgctttca	10500
gtagttcagc ttatttatgc ccactagaaa catcttgtag aagctgaaca ctggggctcc	10560
agattagtgg taaaacctac ttatacaat catagaatca tagaatggcc tgggttgga	10620
gggaccccaa ggatcatgaa gatccaacac ccccgccaca ggcagggcca ccaacctcca	10680
gatctggtac tagaccaggc agcccagggc tccatccaac ctggccatga acacctccag	10740
ggatggagca tccacaacct ctctgggcag cctgtgccag cacctacca ccctctctgt	10800
gaagaacttt tccctgacat ccaatctaag ccttcctcc ttgaggtag atccactccc	10860
ccttggtgcta tcaactgtcta ctctgtaaa aagttgattc tcctcctttt tggaaggttg	10920
caatgaggtc tccttgtagc ctcttctct tctgcaggat gaacaagccc agctccctca	10980
gcctgtcttt ataggagagg tgctccagcc ctctgatcat ctttgtggcc ctctctgga	11040
cccgtccaa gagctccaca tctttcctgt actgggggccc ccaggcctga atgcagtact	11100
ccagatgggg cctcaaaaga gcagagtaaa gagggacaat caccttcctc accctgctgg	11160
ccagccctct tctgatggag ccctggatac aactggcttt ctgagctgca acttctcctt	11220
atcagttcca ctattaaaac aggaacaata caacaggtgc tgatggccag tgcagagttt	11280

025CIP4 PCT SEQ List.txt

ttcacacttc	ttcatttcgg	tagatccttag	atgaggaacg	ttgaagtgtg	gcttctgcgt	11340
gtgcttcttc	ctcctcaaat	actcctgcct	gatacctcac	cccacctgcc	actgaatggc	11400
tccatggccc	cctgcagcca	gggccctgat	gaacccggca	ctgcttcaga	tgctgtttaa	11460
tagcacagta	tgaccaagtt	gcacctatga	atacacaac	aatgtgttgc	atccttcagc	11520
acttgagaag	aagagccaaa	tttgcatgtg	caggaaatgg	tttagtaatt	ctgccaatta	11580
aaacttgttt	atctaccatg	gctgttttta	tggctgttag	tagtggtaca	ctgatgatga	11640
acaatggcta	tgagtaaaaa	tcaagactgt	agatattgca	acagactata	aaattcctct	11700
gtggccttagc	caatgtggta	cttcccacat	tgtataagaa	atttggaag	tttagagcaa	11760
tgtttgaagt	gttgggaaat	ttctgtatac	tcaagagggc	gtttttgaca	actgtagaac	11820
agaggaatca	aaagggggtg	ggaggaagtt	aaaagaagag	gcaggtgcaa	gagagcttgc	11880
agtcccgcgtg	tgtgtacgac	actggcaaca	tgaggtcttt	gctaattcttg	gtgctttgct	11940
tcctgcccct	ggctgcctta	gggtgcgatc	tgctcagac	ccacagcctg	ggcagcagga	12000
ggaccctgat	gctgctggct	cagatgagga	gaatcagcct	gttttagctgc	ctgaaggata	12060
ggcacgattt	tggctttcct	caagaggagt	ttggcaacca	gtttcagaag	gctgagacca	12120
tcctgtgctg	gcacgagatg	atccagcaga	tctttaacct	gttttagcacc	aaggatagca	12180
gcgctgcttg	ggatgagacc	ctgctggata	agttttacac	cgagctgtac	cagcagctga	12240
acgatctgga	ggcttgcgtg	atccagggcg	tgggcgtgac	cgagaccctt	ctgatgaagg	12300
aggatagcat	cctggctgtg	aggaagtact	ttcagaggat	caccctgtac	ctgaaggaga	12360
agaagtacag	cccctgcgct	tgggaagtgc	tgagggtgta	gatcatgagg	agcttttagcc	12420
tgagaccaa	cctgcaagag	agcttgaggt	ctaaggagta	aaaagtctag	agtcggggcg	12480
gccggccgct	tcgagcagac	atgataagat	acattgatga	gtttggacaa	accacaacta	12540
gaatgcagtg	aaaaaaatgc	tttattttgtg	aaattttgtga	tgctattgct	ttattttgtaa	12600
ccattataag	ctgcaataaa	caagttaaca	acaacaattg	cattcatttt	atgtttcagg	12660
ttcaggggga	ggtgtgggag	gtttttttaa	gcaagtaaaa	cctctacaaa	tgtggtaaaa	12720
tcgataagga	tccgtcgacc	gatgcccttg	agagccttca	accagtcag	ctccttccgg	12780
tgggcgcggg	gcatgactat	cgtcgccgca	cttatgactg	tcttctttat	catgcaactc	12840
gtaggacagg	tgccggcagc	gctcttccgc	ttcctcgctc	actgactcgc	tgcgctcggt	12900
cgttcggctg	cggcgagcgg	tatcagctca	ctcaaaggcg	gtaatacggt	tatccacaga	12960
atcaggggat	aacgcaggaa	agaacatgtg	agcaaaaggc	cagcaaaagg	ccaggaaccg	13020
taaaaaggcc	gcgttgctgg	cgtttttcca	taggctccgc	ccccctgacg	agcatcacia	13080
aaatcgacgc	tcaagtcaga	ggtggcgaaa	cccgacagga	ctataaagat	accaggcggt	13140

025CIP4 PCT SEQ List.txt

tccccctgga agctccctcg tgcgctctcc tgttccgacc ctgccgctta ccggatacct	13200
gtccgccttt ctcccttcgg gaagcgtggc gctttctcaa tgctcacgct gtaggtatct	13260
cagttcggtg taggtcgttc gctccaagct gggctgtgtg cacgaacccc ccgttcagcc	13320
cgaccgctgc gccttatccg gtaactatcg tcttgagtcc aaccgggtaa gacacgactt	13380
atcgccactg gcagcagcca ctggtaacag gattagcaga gcgaggtatg taggcggtgc	13440
tacagagttc ttgaagtggg ggcctaacta cggctacact agaaggacag tatttggtat	13500
ctgcgctctg ctgaagccag ttaccttcgg aaaaagagtt ggtagctctt gatccggcaa	13560
acaaaccacc gctggtagcg gtggtttttt tgtttgcaag cagcagatta cgcgcagaaa	13620
aaaaggatct caagaagatc ctttgatctt ttctacgggg tctgacgctc agtggaacga	13680
aaactcacgt taagggatth ttggtcatgag attatcaaaa aggatcttca cctagatcct	13740
tttaaattaa aaatgaagtt ttaaataaat cttaaagtata tatgagtaaa cttgggtctga	13800
cagttaccaa tgcttaataca gtgaggcacc tatctcagcg atctgtctat ttcgttcac	13860
catagttgcc tgactccccg tcgtgtagat aactacgata cgggaggggct taccatctgg	13920
ccccagtgtc gcaatgatac cgcgagaccc acgctcaccg gctccagatt tatcagcaat	13980
aaaccagcca gccggaaggg ccgagcgag aagtggctct gcaactttat ccgcctccat	14040
ccagtctatt aattgttgcc gggaagctag agtaagtagt tcgccagtta atagtttgcg	14100
caacgttggt gccattgcta caggcatcgt ggtgtcacgc tcgtcgtttg gtatggcttc	14160
attcagctcc gggtcccaac gatcaaggcg agttacatga tccccatgt tgtgcaaaaa	14220
agcggttagc tccttcgggtc ctccgatcgt tgtcagaagt aagttggccg cagtgttatc	14280
actcatgggt atggcagcac tgcataattc tcttactgtc atgccatccg taagatgctt	14340
ttctgtgact ggtgagtact caaccaagtc attctgagaa tagtgtagtc ggcgaccgag	14400
ttgtctttgc ccggcgtcaa tacgggataa taccgcgcca catagcagaa ctttaaaagt	14460
gctcatcatt ggaaaacgtt cttcggggcg aaaactctca aggatcttac cgctgttgag	14520
atccagttcg atgtaacca ctctgtcacc caactgatct tcagcatctt ttactttcac	14580
cagcgtttct ggggtgagcaa aaacaggaag gcaaaatgcc gcaaaaaagg gaataagggc	14640
gacacggaaa tggtgaatac tcatactctt ctttttcaa tattattgaa gcatttatca	14700
gggttattgt ctcatgagcg gatacatatt tgaatgtatt tagaaaaata aacaaatagg	14760
ggttccgcgc acatttcccc gaaaagtgcc acctgacgcg ccctgtagcg gcgcattaag	14820
cgcggcgggt gtggtggtta cgcgacgctg gaccgctaca cttgccagcg ccctagcgcc	14880
cgctcctttc gctttcttcc cttcctttct cgccacgttc gccggctttc cccgtcaagc	14940
tctaaatcgg gggctccctt tagggttccg atttagtgct ttacggcacc tcgaccccaa	15000
aaaacttgat taggggtgat gtacacgtag tgggccatcg ccctgataga cggtttttcg	15060

O25CIP4 PCT SEQ List.txt

ccctttgacg	ttggagtcca	cgttcttttaa	tagtggactc	ttgttcctaaa	ctggaacaac	15120
actcaaccct	atctcggctc	attctttttga	tttataaggg	atcttgccga	tttcggccta	15180
ttggttaaaa	aatgagctga	tttaacaaaa	atttaacgcg	aatttttaaca	aaatattaac	15240
gtttacaatt	tcccatctgc	cattcaggct	gcgcaactgt	tgggaagggc	gatcgggtgcg	15300
ggcctcttcg	ctattacgcc	agcccaagct	accatgataa	gtaagtaata	ttaaggtacg	15360
ggaggtactt	ggagcggccg	ctctagaact	agtggatccc	cgggccgcaa	taaaatatct	15420
ttattttcat	tacatctgtg	tgttggtttt	ttgtgtgaat	cgatagtact	aacatacgct	15480
ctccatcaaa	acaaaacgaa	acaaaacaaa	ctagcaaaat	aggctgtccc	cagtgcgaagt	15540
gcaggtgcca	gaacatttct	ctatcgatag	gtaccgagct	cttacgcgtg	ctagccctcg	15600
agcaggatct	atacattgaa	tcaatattgg	caattagcca	tattagtcac	tggttatata	15660
gcataaatca	atattggcta	ttggccattg	catacgttgt	atctatatca	taatattgtac	15720
atttatattg	gctcatgtcc	aatatgaccg	ccatgttgac	attgattatt	gactagttat	15780
taatagtaat	caattacggg	gtcattagtt	catagcccat	atatggagtt	ccgcgttaca	15840
taacttacgg	taaatggccc	gcctggctga	ccgcccacg	accccgccc	attgacgtca	15900
ataatgacgt	atgttcccat	agtaacgcca	atagggactt	tccattgacg	tcaatgggtg	15960
gagtatttac	ggtaaactgc	ccacttgcca	gtacatcaag	tgtatcatat	gccaagtccg	16020
ccccctattg	acgtcaatga	cggtaaatgg	cccgcctggc	attatgcca	gtacatgacc	16080
ttacgggact	ttcctacttg	gcagtacatc	tacgtattag	tcacgcctat	taccatgggtg	16140
atgcggtttt	ggcagtacat	caatgggctg	ggatagcggg	ttgactcacg	gggattttcca	16200
agtctccacc	ccattgacgt	caatgggagt	ttgttttggc	accaaataca	acgggacttt	16260
ccaaaatgtc	gtaacaactc	cgccccattg	acgcaaattg	gcggtaggcg	tgtacgggtg	16320
gagggtctata	taagcagagc	tcgttttagtg	aaccgtcaga	tcgcctggag	acgccatcca	16380
cgctgttttg	acctccatag	aagacaccgg	gaccgatcca	gcctcccctc	gaagctcgac	16440
tctaggggct	cgagatctgc	gatctaagta	agcttgcatg	cctgcaggtc	ggccgccacg	16500
accggtgccg	ccaccatccc	ctgaccacg	cccctgaccc	ctcacaagga	gacgaccttc	16560
catgaccgag	tacaagccca	cggtgcgcct	cgccaccgcg	gacgacgtcc	cccgggccgt	16620
acgcaccctc	gccgccgcgt	tcgccgacta	ccccgccacg	cgccacaccg	tcgaccggga	16680
ccgccacatc	gagcgggtca	ccgagctgca	agaactcttc	ctcacgcgcg	tcgggctcga	16740
catcggcaag	gtgtgggtcg	cggacgacgg	cgccgcgggtg	gcgggtctgga	ccacgccggga	16800
gagcgtcgaa	gcggggggcg	tgctcgccga	gatcggcccc	cgcatggccg	agttgagcgg	16860
ttcccggtcg	gccgcgcagc	aacagatgga	aggcctcctg	gcgccgcacc	ggcccaagga	16920

025CIP4 PCT SEQ List.txt

gcccgcgtgg ttcttgccca ccgtcggcgt ctgcgccgac caccagggca agggctctggg 16980
 cagcgccgtc gtgctccccg gagtggaggc ggccgagcgc gccggggtgc ccgccttcct 17040
 ggagacctcc gcgccccgca acctccccct ctacgagcgg ctcggtttca ccgtcaccgc 17100
 cgacgtcgag gtgcccgaag gaccgcgcac ctggtgcatg acccgcaagc ccggtgcctg 17160
 acgcccggcc cagcaccgc agcggccgac cgaaaggagc gcacgacccc atgggtccga 17220
 ccgaagccga cccgggaggc cccggccgacc ccgcaccgc ccccgaggcc caccgactct 17280
 agagtcgggg cggccggccg cttcgagcag acatgataag atacattgat gagtttggac 17340
 aaaccacaac tagaatgcag tgaaaaaaat gctttatttg tgaaatttgt gatgctattg 17400
 ctttatttgt aaccattata agctgcaata aacaagttaa caacaacaat tgcattcatt 17460
 ttatgtttca ggttcagggg gaggtgtggg aggtttttta aagcaagtaa aacctctaca 17520
 aatgtggtaa aatcgataag gatcaattcg gcttcaggta ccgtcgacga ttaggtcac 17580
 ggtctcgaag ccgcggtgcg ggtgccaggc cgtgcccttg ggctccccgg gcgcgtactc 17640
 cacctacccc atctggtcca tcatgatgaa cgggtcgagg tggcggtagt tgatcccggc 17700
 gaacgcgcgg cgcaccggga agccctcgcc ctcgaaaccg ctgggcgcgg tggtcacggg 17760
 gagcacggga cgtgcgacgg cgtcggcggg tgcggatacg cggggcagcg tcagcgggtt 17820
 ctgcaggtc acggcgggca tgtcgacagc cgaattgatc cgtcgaccga tgcccttgag 17880
 agccttcaac ccagtcagct ccttccgggt ggcgcggggc atgactatcg tcgccgact 17940
 tatgactgtc ttctttatca tgcaactcgt aggacagggt ccggcagcgc tcttccgctt 18000
 cctcgctcac tgactcgctg cgctcggctg ttcggctgcg gcgagcggta tcagctcact 18060
 caaaggcggg aatacggtta tccacagaat caggggataa cgcaggaaag aacatg 18116

<210> 8
 <211> 17402
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> Plasmid pOMIFN-Ins-CMV-pur-attB

<400> 8
 ggccgccacc gcggtggagc tccaattcgc cctatagtga gtcgtattac aattcactgg 60
 ccgtcgtttt acaacgtcgt gactgggaaa accctggcgt taccctaactt aatcgccttg 120
 cagcacatcc ccctttcgcc agctggcgta atagcgaaga ggcccgacc gatcgccctt 180
 cccaacagtt gcgcagcctg aatggcgaat gggacgcgcc ctgtagcggc gcattaagcg 240
 cggcgggtgt ggtggttacg cgcagcgtga ccgctacact tgccagcgcc ctagcgcccg 300
 ctcttttcgc tttcttcctt tcctttctcg ccacgttcgc cggctttccc cgtcaagctc 360
 taaatcgggg gtcctcttta gggttccgat ttagtgcttt acggcacctc gaccccaaaa 420

025CIP4 PCT SEQ List.txt

aacttgatta	gggtgatggt	tcacgtagtg	ggccatcgcc	ctgatagacg	gtttttcgcc	480
ctttgacggt	ggagtcacg	ttctttaata	gtggactctt	gttccaaact	ggaacaacac	540
tcaaccctat	ctcggctctat	tcttttgatt	tataagggat	tttgccgatt	tcggcctatt	600
ggttaaaaaa	tgagctgatt	taacaaaaat	ttaacgcgaa	ttttaacaaa	atattaacgc	660
ttacaattta	ggtggcactt	ttcggggaaa	tgtgcgcgga	acccctatct	gtttatcttt	720
ctaaatacat	tcaaatatgt	atccgctcat	gagacaataa	ccctgataaa	tgcttcaata	780
atattgaaaa	aggaagagta	tgagtattca	acatttccgt	gtcgccctta	ttcccttttt	840
tgcggcattt	tgcttccctg	tttttgctca	cccagaaacg	ctggtgaaag	taaaagatgc	900
tgaagatcag	ttgggtgcac	gagtgggtta	catcgaactg	gatctcaaca	gcggtgaagat	960
ccttgagagt	tttcgccccg	aagaacgttt	tccaatgatg	agcactttta	aagtctctgt	1020
atgtggcgcg	gtattatccc	gtattgacgc	cgggcaagag	caactcggtc	gccgcataca	1080
ctattctcag	aatgacttgg	ttgagtactc	accagtcaca	gaaaagcatc	ttacggatgg	1140
catgacagta	agagaattat	gcagtgctgc	cataaccatg	agtgataaca	ctgcggccaa	1200
cttacttctg	acaacgatcg	gaggaccgaa	ggagctaacc	gcttttttgc	acaacatggg	1260
ggatcatgta	actcgccttg	atcggtggga	accggagctg	aatgaagcca	taccaaacga	1320
cgagcgtgac	accacgatgc	ctgtagcaat	ggcaacaacg	ttgcgcaaac	tattaactgg	1380
cgaactactt	actctagctt	cccggcaaca	attaatagac	tggatggagg	cggataaagt	1440
tgcaggacca	cttctgcgct	cggcccttcc	ggctggctgg	tttattgctg	ataaatctgg	1500
agccggtgag	cgtgggtctc	gcggtatcat	tgagcactg	gggccagatg	gtaagccctc	1560
ccgtatcgta	gttatctaca	cgacggggag	tcaggcaact	atggatgaac	gaaatagaca	1620
gatcgctgag	ataggtgcct	cactgattaa	gcattggtaa	ctgtcagacc	aagtttactc	1680
atatatactt	tagattgatt	taaaacttca	tttttaattt	aaaaggatct	aggtgaagat	1740
cctttttgat	aatctcatga	ccaaaatccc	ttaacgtgag	ttttcgttcc	actgagcgtc	1800
agaccccgta	gaaaagatca	aaggatcttc	ttgagatcct	ttttttctgc	gcgtaatctg	1860
ctgcttgcaa	acaaaaaac	caccgctacc	agcggtggtt	tgtttgccgg	atcaagagct	1920
accaactctt	tttccgaagg	taactggctt	cagcagagcg	cagataccaa	atactgtcct	1980
tctagtgtag	ccgtagttag	gccaccactt	caagaactct	gtagcaccgc	ctacatacct	2040
cgctctgcta	atcctgttac	cagtggctgc	tgccagtggc	gataagtcgt	gtcttaccgg	2100
gttggactca	agacgatagt	taccggataa	ggcgcagcgg	tcgggctgaa	cgggggggttc	2160
gtgcacacag	cccagcttgg	agcgaacgac	ctacaccgaa	ctgagatacc	tacagcgtga	2220
gctatgagaa	agcgccacgc	ttcccgaagg	gagaaaggcg	gacaggatct	cggtaagcgg	2280

025CIP4 PCT SEQ List.txt

caggggtcggg	acaggagagc	gcacgagggg	gcttccaggg	ggaaacgcct	ggtatcttta	2340
tagtcctgtc	gggtttcgcc	acctctgact	tgagcgtcga	tttttgtgat	gctcgtcagg	2400
ggggcgagc	ctatggaaaa	acgccagcaa	cgcggccttt	ttacggttcc	tggccttttg	2460
ctggcctttt	gctcacatgt	tctttcctgc	gttatcccct	gattctgtgg	ataaccgtat	2520
taccgccttt	gagtgagctg	ataccgctcg	ccgcagccga	acgaccgagc	gcagcgagtc	2580
agtgagcgag	gaagcggaag	agcgcccaat	acgcaaaccg	cctctccccg	cgcgttggcc	2640
gattcattaa	tgcagctggc	acgacagggt	tcccgactgg	aaagcgggca	gtgagcgcaa	2700
cgcaattaat	gtgagttagc	tcactcatta	ggcaccccag	gctttacact	ttatgcttcc	2760
ggctcgtatg	ttgtgtggaa	ttgtgagcgg	ataacaattt	cacacaggaa	acagctatga	2820
ccatgattac	gccaagctcg	aaattaaccc	tcactaaagg	gaacaaaagc	tgggtaccgg	2880
gccccccctc	gactagaggg	acagcccccc	cccaaagccc	ccagggatgt	aattacgtcc	2940
ctcccccgct	agggggcagc	agcgagccgc	ccggggctcc	gctccgggtc	ggcgctcccc	3000
ccgcatcccc	gagccggcag	cgtgcgggga	cagcccgggc	acggggaagg	tggcacggga	3060
tcgctttcct	ctgaacgctt	ctcgctgctc	tttgagcctg	cagacacctg	gggggatacg	3120
gggaaaaaagc	tttaggctga	aagagagatt	tagaatgaca	gaatcataga	acggcctggg	3180
ttgcaaagga	gcacagtgtc	catccagatc	caaccccctg	ctatgtgcag	ggtcatcaac	3240
cagcagccca	ggctgcccag	agccacatcc	agcctggcct	tgaatgcctg	cagggatggg	3300
gcatccacag	cctccttggg	caacctgttc	agtgcgtcac	caccctctgg	gggaaaaact	3360
gcctcctcat	atccaacca	aacctcccct	gtctcagtgt	aaagccattc	ccccttgtcc	3420
tatcaagggg	gagtttgctg	tgacattgtt	ggtctggggg	gacacatgtt	tgccaattca	3480
gtgcatcacg	gagaggcaga	tcttggggat	aaggaaagtgc	aggacagcat	ggacgtggga	3540
catgcagggtg	ttgagggctc	tgggacactc	tccaagtcac	agcgttcaga	acagccttaa	3600
ggataagaag	ataggataga	aggacaaaga	gcaagttaaa	accagcatg	gagaggagca	3660
caaaaaggcc	acagacactg	ctggtccttg	tgtctgagcc	tgcatgtttg	atgggtgtctg	3720
gatgcaagca	gaaggggtgg	aagagcttgc	ctggagagat	acagctgggt	cagtaggact	3780
gggacaggca	gctggagaat	tgccatgtag	atgttcatac	aatcgtcaaa	tcatgaaggc	3840
tggaaaagcc	ctccaagatc	cccaagacca	accccaaccc	accacaccgtg	cccactggcc	3900
atgtccctca	gtgccacatc	cccacagttc	ttcatcacct	ccagggacgg	tgaccccccc	3960
acctccgtgg	gcagctgtgc	cactgcagca	ccgctctttg	gagaaggtaa	atcttgctaa	4020
atccagccccg	accctcccct	ggcacaacgt	aaggccatta	tctctcatcc	aactccagga	4080
cggagtcagt	gaggatgggg	ctctagtcga	ggtcgacggg	atcgataagc	ttgattaggc	4140
agagcaatag	gactctcaac	ctcgtgagta	tggcagcatg	ttaactctgc	actggagtcc	4200

025CIP4 PCT SEQ List.txt

agcgtgggaa	acaatctgcc	ttgcacatga	gtcttcgtgg	gccaatattc	cccaacggtt	4260	
ttccttcagc	ttgtcttgtc	tcctaagctc	tcaaaacacc	tttttggtga	ataaactcac	4320	
ttggcaacgt	ttatctgtct	taccttagtg	tcacgtttca	tcctattcc	cctttctcct	4380	
cctccgtgtg	gtacacagtg	gtgcacactg	gttcttctgt	tgatgttctg	ctctgacagc	4440	
caatgtgggt	aaagtcttc	ctgccacgtg	tctgtgttgt	tttcacttca	aaaagggccc	4500	
tgggctcccc	ttggagctct	caggcatctc	cttaatcatc	acagtcacgc	tggcaggatt	4560	
agtccctcct	aaaccttaga	atgacctgaa	cgtgtgctcc	ctctttgtag	tcagtgcagg	4620	
gagacgtttg	cctcaagatc	agggctccatc	tcacccacag	ggccattccc	aagatgaggt	4680	
ggatggttta	ctctcacaaa	aagttttctt	atgtttggct	agaaaggaga	actcactgcc	4740	
tacctgtgaa	ttcccctagt	cctggttctg	ctgccactgc	tgctgtgca	gcctgtccca	4800	
tggagggggc	agcaactgct	gtcacaaagg	tgatcccacc	ctgtctccac	tgaaatgacc	4860	
tcagtgccac	gtgttgata	gggtataaag	tacgggaggg	ggatgcccgg	ctcccttcag	4920	
ggttgacag	cagaagtgtc	tgtgtataga	gtgtgtctta	atctattaat	gtaacagaac	4980	
aacttcagtc	ctagtgtttt	gtgggctgga	attgcccatg	tggtagggac	aggcctgcta	5040	
aatcactgca	atcgcttatg	ttctgaaggt	atttgggaaa	gaaagggatt	tgggggattg	5100	
cctgtgattg	gctttaattg	aatggcaa	aatcacagg	aggttctgct	caacagttgg	5160	
ttgtttcagc	caattcttgc	agccaaagag	ccgggtgccc	agcgatataa	tagttgtcac	5220	
ttgtgtctgt	atggatgaca	gggaggtagg	gtgacctgag	gaccaccctc	cagcttctgc	5280	
tagcgtaggt	acagtcacca	cctccagctc	cacacgagtc	ccatcgtggg	ttaccaaaga	5340	
aacacaatta	tttggaccag	tttggaaagt	cacccgctga	attgtgaggc	tagattaata	5400	
gagctgaaga	gcaa	aatgttcc	ccaacttgg	gatactagtt	gggtattagta	tcagaggaac	5460
agggccatag	cacctccatg	ctattagatt	ccggctggca	tgtacttttc	aagatgattt	5520	
gtaactaaca	atggcttatt	gtgcttgtct	taagtctgtg	tccta	aatgta	aatgttcctt	5580
tggtttatat	aaccttcttg	ccatttgctc	ttcaggtgtt	cttgcagaac	actggctgct	5640	
ttaatctagt	ttactgttg	cttgattatt	cttagggata	agatctgaat	aaactttttg	5700	
tggctttggc	agacttttagc	ttgggcttag	ctccacatt	agcttttgct	gccttttctg	5760	
tgaagctatc	aagatcctac	tcaatgacat	tagctgggtg	caggtgtacc	aaatcctgct	5820	
ctgtggaaca	cattgtctga	tgataccgaa	ggcaa	acgtg	aactcaaaga	ggcacagagt	5880
taagaagaag	tctgtgcaat	tcagaggaaa	agccaaagtg	gccattagac	acactttcca	5940	
tgcagcattt	gccagtaggt	ttcatataaa	actacaaaat	ggaataaacc	actacaaatg	6000	
ggaaaagcct	gatactagaa	tttaa	aatatt	cacccaggct	caaggggtgt	ttcatggagt	6060

025CIP4 PCT SEQ List.txt

aatatcactc tataaaagta gggcagccaa ttattcacag acaaagcttt tttttttctg	6120
tgctgcagtg ctgtttttctg gctgatccag gggtacttat tgtgggtctg agagctgaat	6180
gattttctcct tgtgtcatgt tgggtgaagga gatatggcca gggggagatg agcatgttca	6240
agaggaaacg ttgcattttg gtggcttggg agaaaggtag aacgatataca ggtccatagt	6300
gtcactaaga gatctgaagg atggttttac agaacagttg acttggtctg gtgcaggctt	6360
ggctgtaaata ggatggaagg atggacagat ggggtggacag agattttctgt gcaggagatc	6420
atctcctgag ctcggtgctt gacagactgc agatccatcc cataaccttc tccagcatga	6480
gagcgcgggg agcttttggtta ctgttcagtc tgctgcttgt tgcttcctgg gtgcacagtg	6540
gtgatttttct tactcacaca gggcaaaaac ctgagcagct tcaaagtga caggttgctc	6600
tcataggcca ttcagttgtc aagatgaggt ttttggtttc ttgttttgta aggtgggaag	6660
aagcactgaa ggatcagttg cgagggcagg ggttttagcac tgttcagaga agtcttattt	6720
taactcctct catgaacaaa aagagatgca ggtgcagatt ctggcaagca tgcagtgaag	6780
gagaaagccc tgaatttctg atatatgtgc aatgttgggc acctaacatt ccccgctgaa	6840
gcacagcagc tccagctcca tgcagtactc acagctggtg cagccctcgg ctccagggtc	6900
tgagcagtgct tgggactcac gaggttccat gtctttcaca ctgataatgg tccaatttct	6960
ggaatgggtg cccatccttg gaggtcccca aggccaggct ggctgcgtct ccgagcagcc	7020
cgatctggtg gtgagtagcc agcccatggc aggagttaga gcctgatggt ctttaaggtc	7080
ccttccaacc taagccatcc tacgattcta ggaatcatga cttgtgagt tgtattgcag	7140
aggcaatatt ttaaagttat aaatgttttc tccccctcct tgtttgcata agttatcttg	7200
atcgccttat caatgctttt ggagtctcca gtcatttttc ttacamcaaa aagaggagga	7260
agaatgaaga gaatcattta atttcttgat tgaatagtag gattcagaaa gctgtacgta	7320
atgccgtctc tttgtatcga gctgtaaggt ttctcatcat ttatcagcgt ggtacatatc	7380
agcacttttc catctgatgt ggaaaaaaaa atccttatca tctacagtct ctgtacctaa	7440
acatcgctca gactctttac caaaaaagct ataggtttta aaactacatc tgctgataat	7500
ttgccttggt ttagctcttc ttccatatgc tgcgtttgtg agagggtcgt ggatgggcct	7560
aaactctcag ctgctgagct tgatgggtgc ttaagaatga agcactcact gctgaaactg	7620
ttttcatttc acaggaatgt tttagtggca ttgtttttat aactacatat tcctcagata	7680
aatgaaatcc agaaataatt atgcaaactc actgcatccg ttgcacaggt ctttatctgc	7740
tagcaaagga aataatttgg ggatggcaaa aacattcctt cagacatcta tatttaaagg	7800
aatataatcc tgggtaccac ccacttcatc cctcattatg ttcacactca gagatactca	7860
ttctcttggt gttatcattt gatagcgttt tctttggttc tttgccacgc tctgggctat	7920
ggctgcacgc tctgcactga tcagcaagta gatgcgaggg aagcagcagt gagaggggct	7980

025CIP4 PCT SEQ List.txt

gccctcagct	ggcaccagc	cgctcagcct	aggaggggac	cttgcctttc	caccagctga	8040
ggtgcagccc	tacaagctta	cacgtgctgc	gagcaggtga	gcaaagggag	tcttcatggt	8100
gtgtttcttg	ctgcccggaa	gcaaaacttt	actttcattc	attccccttg	aagaatgagg	8160
aatgtttgga	aacggactgc	tttacgttca	atttctctct	tccctttaag	gctcagccag	8220
gggccattgc	tgaggacggc	atcggggccc	cctggaccaa	atctgtggca	cagatggttt	8280
cacttacatc	agtggatgtg	ggatctgctc	ctgtaatgtg	tccttctgaa	ggaaggaacg	8340
tgctttccaa	gtgccagccc	cacagccccc	agccccctcc	tgtgctgctc	caattcatct	8400
cctcttcctc	cttctccctt	tgctgtttgt	gctcgggtag	aaatcatgaa	gatttagaag	8460
agaaaacaaa	ataactggag	tggaaaccca	ggtgatgcag	ttcattcagc	tgtcataggt	8520
ttgtcgttgc	tataggtctg	tatcagagat	gctarcacca	ctttgctgtc	ggtgcttaac	8580
tcgggtgaac	tctccttcac	tcgcatcatt	tgcgggcctt	atttacatcc	ccagcatcca	8640
tcaccctctg	ggaaaatggg	cgactggat	ctctaattga	agactttccc	tctttcagag	8700
cctgtgggat	gtgcagtgac	aagaaacgtg	gaggggctga	gcagcagcac	tgccccagg	8760
gagcaggagc	ggatgccatc	ggtggcagca	tcccaaata	tgtagcgga	tgctgagcag	8820
gcagcggacg	aacggacaga	agcgatgcgt	acaccttctg	ttgacatggt	atttggcagc	8880
gatttaacac	tcgcttccta	gtcctgctat	tctccacagg	ctgcattcaa	atgaacgaag	8940
ggaagggagg	caaaaagatg	caaaatccga	gacaagcagc	agaaatattt	cttcgctacg	9000
gaagcgtgcg	caaacaacct	tctccaacag	caccagaaga	gcacagcgta	acctttttca	9060
agaccagaaa	aggaaattca	caaagcctct	gtggatacca	gcgcgttcag	ctctcctgat	9120
agcagatttc	ttgtcagggt	gcgaatggg	tatggtgcc	ggaggtgcag	ggaccatatg	9180
atcatataca	gcacagcagt	cattgtgcat	gtattaatat	atattgagta	gcagtgttac	9240
tttgccaaag	caatagttca	gagatgagtc	ctgctgcata	cctctatctt	aaaactaact	9300
tataaatagt	aaaaccttct	cagttcagcc	acgtgctcct	ctctgtcagc	accaatggtg	9360
cttcgcctgc	accagctgc	aaggaatcag	ccggtgatct	cattaacact	cagctctgca	9420
ggataaatta	gattgtttcca	ctctcttttg	ttgttaatta	cgacggaaca	attgttcagt	9480
gctgatggtc	ctaattgtca	gctacagaaa	acgtctccat	gcagttcctt	ctgcgccagc	9540
aaactgtcca	ggctatagca	ccgtgatgca	tgctacctct	cactccatcc	ttcttctctt	9600
tcccaccagg	gagagctgtg	tgttttcact	ctcagccact	ctgaacaata	ccaaactgct	9660
acgcactgcc	tccctcggaa	agagaatccc	cttggttgctt	ttttattttac	aggatccttc	9720
ttaaaaagca	gaccatcatt	cactgcaaac	ccagagcttc	atgcctctcc	ttccacaacc	9780
gaaaacagcc	ggcttcattt	gtctttttta	aatgctgttt	tccaggtgaa	ttttggccag	9840

025CIP4 PCT SEQ List.txt

cgtgttggct	gagatccagg	agcacgtgtc	agctttctgc	tctcattgct	cctgttctgc	9900
attgcctctt	tctgggggtt	ccaagagggg	gggagacttt	gcgcggggat	gagataatgc	9960
cccttttctt	agggtggctg	ctgggcagca	gagtggctct	gggtcactgt	ggcaccaatg	10020
ggaggcacca	gtgggggtgt	gttttgtgca	ggggggaagc	attcacagaa	tggggctgat	10080
cctgaagctt	gcagtccaag	gctttgtctg	tgtaccagct	gaaatccttc	ctctgttaca	10140
taaagcccag	ataggactca	gaaatgtagt	cattccagcc	cccctcttcc	tcagatctgg	10200
agcagcactt	gtttgcagcc	agtcctcccc	aaaatgcaca	gacctcgccg	agtggaggga	10260
gatgtaaaca	gcgaagggtta	attacctcct	tgtcaaaaac	actttgtggg	ccatagatgt	10320
ttctgtcaat	cttacaaaac	agaaccgaga	ggcagcgagc	actgaagagc	gtgttcccat	10380
gctgagttaa	tgagacttgg	cagctcgctg	tgagagatg	atccctgtgc	ttcatgggag	10440
gctgtaacct	gtctcccat	cgcttccaca	ccgcagtgtc	gtcctggaca	cctcacctc	10500
cataagctgt	aggatgcagc	tgcccaggga	tcaagagact	tttcctaagg	ctcttaggac	10560
tcacttttgc	cgctcagtag	cgtgcagcaa	ttactcatcc	caactatact	gaatgggttt	10620
ctgccagctc	tgcttgtttg	tcaataagca	tttcttcatt	ttgcctctaa	gtttctctca	10680
gcagcaccgc	tctgggtgac	ctgagtggcc	acctggaacc	cgaggggcac	agccaccacc	10740
tccctgttgc	tgctgctcca	gggactcatg	tgctgctgga	tggggggaag	catgaagtgc	10800
ctcaccacga	cacctgggtt	gcaatggctg	cagcgtgtc	ttcttgggat	gcagattgtt	10860
tccagccatt	actttagaaa	atgtgtgtgt	gaagcccttt	gtatctcttt	ctgtggccct	10920
tcagcaaaaag	ctgtgggaaa	gctctgaggc	tgctttcttg	ggcgtgggag	gaattgtatg	10980
ttccttcttt	aacaaaaatt	atccttagga	gagagcactg	tgcaagcatt	gtgcacataa	11040
aacaattcag	gttgaaaggg	ctctctggag	gtttccagcc	tgactactgc	tcgaagcaag	11100
gccagggttca	aagatgggtc	aggatgtgtg	gtgccttcct	gattatctgt	gccaccaatg	11160
gaggagattc	acagccactc	tgcttcccgt	gccactcatg	gagaggaata	ttcccttata	11220
ttcagataga	atgttatcct	ttagctcagc	cttccctata	accccatgag	ggagctgcag	11280
atccccatac	tctccccttc	tctgggggtga	aggccgtgtc	cccagcccc	ccttcccacc	11340
ctgtgcccta	agcagcccgc	tgccctctgc	tggtgtgtgt	cctatatgtc	aatgcctgtc	11400
cttgacgtcc	agcctgggac	atttaattca	tcaccagggt	aatgtggaac	tgtgtcatct	11460
tcccctgcag	ggtacaaaagt	tctgcacggg	gtcctttcgg	ttcaggaaaa	ccttcactgg	11520
tgctacctga	atcaagctct	atttaataag	ttcataagca	catggatgtg	ttttcctaga	11580
gatacgtttt	aatgggtatca	gtgattttta	tttgctttgt	tgcttacttc	aaacagtgcc	11640
tttgggcagg	agggtgaggga	cgggtctgcc	gttggctctg	cagtgatttc	tccaggcgtg	11700
tggctcagggt	cagatagtgg	tcactctgtg	gccagaagaa	ggacaaagat	ggaaattgca	11760

025CIP4 PCT SEQ List.txt

gattgagtca	cgtaagcag	gcattcttga	gtgatttgag	gcagtttcat	gaaagagcta	11820
cgaccactta	ttgtttgttt	ccccctttac	aacagaagtt	ttcatcaaaa	taacgtggca	11880
aagcccagga	atgtttggga	aaagtgtagt	taaatgtttt	gtaattcatt	tgtcggagtg	11940
ctaccagcta	agaaaaaagt	cctacctttg	gtatggtagt	cctgcagaga	atacaacatc	12000
aattattagt	tggaaaaaaa	caccaccacc	accagaaact	gtaatggaaa	atgtaaacca	12060
agaaattcct	tgggtaagag	agaaaggatg	tcgtatactg	gccaaagtcct	gcccagctgt	12120
cagcctgctg	accctctgca	gttcaggacc	atgaaacgtg	gcactgtaag	acgtgtcccc	12180
tgcctttgct	tgccacaga	tctctgccct	tgtgctgact	cctgcacaca	agagcatttc	12240
cctgtagcca	aacagcgatt	agccataagc	tgcacctgac	tttgaggatt	aagagtttgc	12300
aattaagtgg	attgcagcag	gagatcagtg	gcagggttgc	agatgaaatc	cttttctagg	12360
ggtagctaag	ggctgagcaa	cctgtcctac	agcacaagcc	aaaccagcca	agggttttcc	12420
tgtgctgttc	acagaggcag	ggccagctgg	agctggagga	ggttgctgctg	ggacccttct	12480
ccctgtgctg	agaatggagt	gatttctggg	tgtgtttcct	gtggcttgca	ctgagcagct	12540
caagggagat	cggtgctcct	catgcagtg	caaaactcgt	gtttgatgca	gaaagatgga	12600
tgtgcacctc	cctcctgcta	atgcagccgt	gagcttatga	aggcaatgag	ccctcagtg	12660
agcaggagct	gtagtgcact	cctgtaggtg	ctagggaaaa	tctctggttc	ccagggatgc	12720
attcataagg	gcaatatatc	ttgaggctgc	gccaaatctt	tctgaaatat	tcatgcgtgt	12780
tcccttaatt	tatagaaaca	aacacagcag	aataattatt	ccaatgcctc	ccctcgaagg	12840
aaaccatat	ttccatgtag	aatgtaacc	tatatacaca	cagccatgct	gcattccttca	12900
gaacgtgcca	gtgctcatct	cccatggcaa	aatactacag	gtattctcac	tatgttggac	12960
ctgtgaaagg	aaccatggta	agaaacttcg	gttaaaggta	tggctgcaaa	actactcata	13020
ccaaaacagc	agagctccag	acctcctctt	aggaaagagc	cacttggaga	gggatgggtgt	13080
gaaggctgga	ggtgagagac	agagcctgtc	ccagttttcc	tgtctctatt	ttctgaaacg	13140
tttgcaggag	gaaaggacaa	ctgtactttc	aggcatagct	ggtgccctca	cgtaaataag	13200
ttccccgaac	ttctgtgtca	tttgttctta	agatgctttg	gcagaacact	ttgagtcaat	13260
tcgcttaact	gtgactaggt	ctgtaaataa	gtgctccctg	ctgataaggt	tcaagtgaca	13320
tttttagtgg	tatttgacag	catttacctt	gctttcaagt	cttctaccaa	gctcttctat	13380
acttaagcag	tgaaaccgcc	aagaaaccct	tccttttatc	aagctagtgc	taaataccat	13440
taacttcata	ggttagatac	ggtgctgcca	gcttcacctg	gcagtgggtg	gtcagttctg	13500
ctggtgacaa	agcctccctg	gcctgtgctt	ttacctagag	gtgaatatcc	aagaatgcag	13560
aactgcatgg	aaagcagagc	tgcaggcacg	atgggtgctga	gccttagctg	cttcctgctg	13620

025CIP4 PCT SEQ List.txt

ggagatgtgg	atgcagagac	gaatgaagga	cctgtccctt	actccccctca	gcattctgtg	13680
ctatttaggg	ttctaccaga	gtccttaaga	ggtttttttt	ttttttgggtc	caaaagtctg	13740
tttgtttgg	tttgaccact	gagagcatgt	gacacttggtc	tcaagctatt	aaccaagtgt	13800
ccagccaaaa	tcaattgcct	gggagacgca	gaccattacc	tggagggtcag	gacctcaata	13860
aatattacca	gcctcattgt	gccgctgaca	gattcagctg	gctgctccgt	gttccagtcc	13920
aacagttcgg	acgccacgtt	tgtatatatt	tgcaggcagc	ctcggggggga	ccatctcagg	13980
agcagagcac	cggcagccgc	ctgcagagcc	gggcagtacc	tcaccatggc	tttgaccttt	14040
gccttactgg	tggctctcct	ggtgctgagc	tgcaagagca	gctgctctgt	gggctgcgat	14100
ctgcctcaga	cccacagcct	gggcagcagg	aggacctga	tgctgctggc	tcagatgagg	14160
agaatcagcc	tgttttagctg	cctgaaggat	aggcacgatt	ttggctttcc	tcaagaggag	14220
tttggcaacc	agtttcagaa	ggctgagacc	atccctgtgc	tgcacgagat	gatccagcag	14280
atctttaacc	tgttttagcac	caaggatagc	agcgtgctt	gggatgagac	cctgctggat	14340
aagttttaca	ccgagctgta	ccagcagctg	aacgatctgg	aggcttgctg	gatccagggc	14400
gtgggcgtga	ccgagacccc	tctgatgaag	gaggatagca	tcctggctgt	gaggaagtac	14460
tttcagagga	tcacctgtga	cctgaaggag	aagaagtaca	gcccctgcgc	ttgggaagtc	14520
gtgagggctg	agatcatgag	gagcttttagc	ctgagcacca	acctgcaaga	gagcttgagg	14580
tctaaggagt	aaaaagtcta	gagtcggggc	ggccggccgc	ttcgagcaga	catgataaga	14640
tacattgatg	agtttgagca	aaccacaact	agaatgcagt	gaaaaaaaaatg	ctttattttgt	14700
gaaatttggtg	atgctattgc	tttattttgta	accattataa	gctgcaataa	acaagttaac	14760
aacaacaatt	gcattcattt	tatgtttcag	gttcaggggg	aggtgtggga	ggtttttttaa	14820
agcaagtaaa	acctctacaa	atgtggtaaa	atcgataccg	tcgacctcga	ctagagcggc	14880
cactaacata	cgctctccat	caaaacaaaa	cgaacaaaaa	caaactagca	aaataggctg	14940
tccccagtg	aagtgcaggt	gccagaacat	ttctctatcg	ataggtaccg	agctcttacg	15000
cgtgctagcc	ctcgagcagg	atctatacat	tgaatcaata	ttggcaatta	gccatattag	15060
tcattgggta	tatagcataa	atcaatattg	gctattggcc	attgcatacg	ttgtatctat	15120
atcataatat	gtacatttat	attggctcat	gtccaatatg	accgcatgt	tgacattgat	15180
tattgactag	ttattaatag	taatcaatta	cggggtcatt	agttcatagc	ccatatatgg	15240
agttccgcgt	tacataactt	acggtaaatg	gcccgcctgg	ctgaccgccc	aacgaccccc	15300
gcccattgac	gtcaataatg	acgtatgttc	ccatagtaac	gccaataggg	actttccatt	15360
gacgtcaatg	ggtggagtat	ttacggtaaa	ctgcccactt	ggcagtacat	caagtgtatc	15420
atatgccaa	tccgccccct	attgacgtca	atgacggtaa	atggcccgc	tggcattatg	15480
cccagtacat	gaccttacgg	gactttccta	cttggcagta	catctacgta	ttagtcatcg	15540

O25CIP4 PCT SEQ List.txt

ctattaccat	ggtgatgcgg	ttttggcagt	acatcaatgg	gcgtggatag	cggtttgact	15600
cacggggatt	tccaagtctc	caccccattg	acgtcaatgg	gagtttgttt	tggcaccaaa	15660
atcaacggga	ctttccaaaa	tgtcgtaca	actccgcccc	attgacgcaa	atgggcggta	15720
ggcgtgtacg	gtgggaggtc	tatataagca	gagctcgttt	agtgaaccgt	cagatgcctt	15780
ggagacgcca	tccacgctgt	tttgacctcc	atagaagaca	ccgggaccga	tccagcctcc	15840
cctcgaagct	cgactctagg	ggctcgagat	ctgcgatcta	agtaagcttg	catgcctgca	15900
ggtcggccgc	cacgaccggt	gccgccacca	tcccctgacc	cacgcccctg	acccttcaca	15960
aggagacgac	cttccatgac	cgagtacaag	cccacggtgc	gcctcgccac	ccgcgacgac	16020
gtcccccg	ccgtacgcac	cctcgccgcc	gcgttcgccg	actacccgc	cacgcgccac	16080
accgtcgacc	cggaccgcca	catcgagcgg	gtcaccgagc	tgcaagaact	cttcctcacg	16140
cgcgtcgggc	tcgacatcgg	caagggtgtg	gtcgcggacg	acggcgccgc	ggtggcggtc	16200
tggaccacgc	cggagagcgt	cgaagcgggg	gcggtgttcg	ccgagatcgg	cccgcgcatg	16260
gccgagttga	gcggttccc	gctggccgcg	cagcaacaga	tggaaggcct	cctggcgccg	16320
caccggccca	aggagcccgc	gtggttcctg	gccaccgtcg	gcgtctcgcc	cgaccaccag	16380
ggcaagggtc	tgggcagcgc	cgtcgtgctc	cccgagtg	aggcgccga	gcgcgccggg	16440
gtgcccgcct	tcctggagac	ctccgcgcc	cgcaacctcc	ccttctacga	gcggctcggc	16500
ttcaccgtca	ccgccgacgt	cgaggtgccc	gaaggaccgc	gcacctggtg	catgaccgc	16560
aagcccgggtg	cctgacgccc	gccccacgac	ccgcagcgcc	cgaccgaaag	gagcgcacga	16620
ccccatggct	ccgaccgaag	ccgaccggg	cggccccgcc	gaccccgcac	ccgccccga	16680
ggcccaccga	ctctagagtc	ggggcgccg	gccgcttcga	gcagacatga	taagatacat	16740
tgatgagttt	ggacaaacca	caactagaat	gcagtgaaa	aaatgcttta	tttgtgaaat	16800
ttgtgatgct	attgctttat	ttgtaaccat	tataagctgc	aataaacaag	ttaacaaca	16860
caattgcatt	cattttatgt	ttcaggttca	gggggaggtg	tgggaggttt	tttaaagcaa	16920
gtaaaacctc	tacaaatgtg	gtaaaatcga	taaggatcaa	ttcggttcga	ggtaccgtcg	16980
acgatgtagg	tcacggtctc	gaagccgcgg	tgcggtgccc	agggcggtcc	cttgggctcc	17040
ccgggcgcgt	actccacctc	acccatctgg	tccatcatga	tgaacgggtc	gaggtggcgg	17100
tagttgatcc	cggcgaacgc	gcggcgacc	gggaagccct	cgccctcgaa	accgctgggc	17160
gcggtggtca	cggtagacac	gggacgtgcg	acggcgtcgg	cgggtgcgga	tacgcggggc	17220
agcgtcagcg	ggttctcgac	ggtcacggcg	ggcatgtcga	cagccgaatt	gatccgtcga	17280
ccgatgccct	tgagagcctt	caaccagtc	agctccttcc	ggtgggcgcg	gggcatgact	17340
atcgtcgccg	cacttatgac	tgtcttcttt	atcatgcaac	tcgtaggaca	ggtgccggca	17400

025CIP4 PCT SEQ List.txt

gc

17402

<210> 9
 <211> 5172
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> Plasmid pRSV-Int

<400> 9
 ctgcattaat gaatcggcca acgcgcgggg agaggcgggt tgcgtattgg gcgctcttcc 60
 gcttcctcgc tcaactgactc gctgcgctcg gtcgttcggc tgcggcgagc ggtatcagct 120
 cactcaaagg cggtaatacg gttatccaca gaatcagggg ataacgcagg aaagaacatg 180
 tgagcaaaaag gccagcaaaa ggccaggaac cgtaaaaagg ccgcgttgct ggcgtttttc 240
 cataggctcc gccccctga cgagcatcac aaaaatcgac gctcaagtca gaggtggcga 300
 aacccgacag gactataaag ataccaggcg tttccccctg gaagctccct cgtgcgctct 360
 cctgttccga ccctgccgt taccggatac ctgtccgcct ttctcccttc ggggaagcgtg 420
 gcgctttctc aatgctcacg ctgtaggtat ctgagttcgg tgtaggtcgt tcgctccaag 480
 ctgggctgtg tgcacgaacc ccccggttcag cccgaccgct gcgccttata cggttaactat 540
 cgtcttgagt ccaacccggt aagacacgac ttatcgccac tggcagcagc cactggtaac 600
 aggattagca gagcgaggta tgtaggcggg gctacagagt tcttgaagtg gtggcctaac 660
 tacggctaca ctagaaggac agtatattgg atctgcgctc tgctgaagcc agttaccttc 720
 ggaaaaagag ttggtagctc ttgatccggc aaacaaacca ccgctggtag cgggtggtttt 780
 tttgtttgca agcagcagat tacgcgcaga aaaaaaggat ctcaagaaga tcctttgatc 840
 ttttctacgg ggtctgacgc tcagtggaaac gaaaactcac gttaagggat tttggtcatg 900
 agattatcaa aaaggatctt cacctagatc cttttaaatt aaaaatgaag ttttaaataca 960
 atctaaagta tatatgagta aacttgggtc gacagttacc aatgcttaat cagtgaggca 1020
 cctatctcag cgatctgtct atttcgttca tccatagttg cctgactccc cgtcgtgtag 1080
 ataactacga tacgggaggg cttaccatct ggccccagtg ctgcaatgat accgcgagac 1140
 ccacgctcac cggctccaga tttatcagca ataaaccagc cagccggaag ggccgagcgc 1200
 agaagtgggtc ctgcaacttt atccgcctcc atccagtcta ttaattgttg ccgggaagct 1260
 agagtaagta gttcgccagt taatagtttg cgcaacggtg ttgccattgc tacaggcatc 1320
 gtggtgtcac gctcgtcgtt tgggtatggc tcattcagct ccggttccca acgatcaagg 1380
 cgagttacat gatcccccat gttgtgcaaa aaagcgggta gtccttcgg tcctccgatac 1440
 gttgtcagaa gtaagtggc cgagtggtta tcaactatgg ttatggcagc actgcataat 1500
 tctcttactg tcatgccatc cgtaagatgc ttttctgtga ctggtgagta ctcaaccaag 1560

025CIP4 PCT SEQ List.txt

tcattctgag	aatagtgtat	gcggcgaccg	agttgctctt	gcccggcgtc	aatacgggat	1620
aataccgcgc	cacatagcag	aactttaaaa	gtgctcatca	ttggaaaacg	ttcttcgggg	1680
cgaaaactct	caaggatctt	accgctgttg	agatccagtt	cgatgtaacc	cactcgtgca	1740
cccaactgat	cttcagcatc	ttttactttc	accagcgttt	ctgggtgagc	aaaaacagga	1800
aggcaaaatg	ccgcaaaaaa	gggaataagg	gcgacacgga	aatgttgaat	actcatactc	1860
ttcctttttc	aatattattg	aagcatttat	cagggttatt	gtctcatgag	cggatacata	1920
tttgaatgta	tttagaaaaa	taaacaaata	ggggttccgc	gcacatttcc	ccgaaaagtg	1980
ccacctgacg	tcgacggatc	gggagatctc	ccgatcccct	atggtcgact	ctcagtacaa	2040
tctgctctga	tgccgcatag	ttaagccagt	atctgctccc	tgcttggtg	ttggaggctg	2100
ctgagtagtg	cgcgagcaaa	atttaagcta	caacaaggca	aggcttgacc	gacaattgca	2160
tgaagaatct	gcttaggggt	aggcgttttg	cgctgcttcg	cgatgtacgg	gccagatata	2220
cgctgctag	gggtctagga	tcgattctag	gaattctcta	gccgcggtct	agggatcccg	2280
gcgcgtatgg	tgactctca	gtacaatctg	ctctgatgcc	gcatagttaa	gccagtatct	2340
gctccctgct	tgtgtgttgg	aggctcgtga	gtagtgcgcg	agcaaaattt	aagctacaac	2400
aaggcaaggc	ttgaccgaca	attgcatgaa	gaatctgctt	agggttaggc	gttttgcgct	2460
gcttcgcgat	gtacgggcca	gatatacgcg	tatctgaggg	gactaggggtg	tgtttagggc	2520
aaaagcgggg	cttcggttgt	acgcggttag	gagtcacctc	aggatatagt	agtttcgctt	2580
ttgcataggg	agggggaaat	gtagtcttat	gcaatacact	tgtagtcttg	caacatggta	2640
acgatgagtt	agcaacatgc	cttacaagga	gagaaaaagc	accgtgcatg	ccgattgggtg	2700
gaagtaagg	ggtacgatcg	tgccctatta	ggaaggcaac	agacaggctc	gacatggatt	2760
ggacgaacca	ctgaattccg	cattgcagag	ataattgtat	ttaagtgcct	agctcgatac	2820
aataaacgcc	atttgaccat	tcaccacatt	ggtgtgcacc	tccaagcttg	catgcctgca	2880
ggtaccggtc	cggaattccc	gggtcgacga	gctcactagt	cgtaggggtcg	ccgacatgac	2940
acaaggggtt	gtgaccgggg	tggacacgta	cgcggttgct	tacgaccgtc	agtcgcgcga	3000
gcgcgagaat	tcgagcgag	caagcccagc	gacacagcgt	agcgccaacg	aagacaaggc	3060
ggccgacctt	cagcgcaag	tcgagcgcg	cgggggccgg	ttcagggttcg	tcgggcattt	3120
cagcgaagcg	ccgggcacgt	cggcgttcgg	gacggcgag	cgcccgaggt	tcgaacgcat	3180
cctgaacgaa	tgccgcgccg	ggcggtctaa	catgatcatt	gtctatgacg	tgtcgcgctt	3240
ctcgcgcctg	aaggctcatg	acgcgattcc	gattgtctcg	gaattgctcg	ccctgggcgt	3300
gacgattgtt	tccactcagg	aaggcgtctt	ccggcaggga	aacgtcatgg	acctgattca	3360
cctgattatg	cggctcgacg	cgtcgcacaa	agaatcttcg	ctgaagtcgg	cgaagattct	3420

025CIP4 PCT SEQ List.txt

cgacacgaag aaccttcagc gcgaattggg cgggtacgtc ggcgggaagg cgccttacgg	3480
cttcagagctt gtttcggaga cgaaggagat cacgcgcaac ggcgaatgg tcaatgtcgt	3540
catcaacaag cttgcgcaact cgaccactcc ccttacccgga cccttcgagt tcgagcccga	3600
cgtaatccgg tgggtgggtggc gtgagatcaa gacgcacaaa caccttccct tcaagccggg	3660
cagtcaagcc gccattcacc cgggcagcat cacggggcctt tgtaagcgca tggacgctga	3720
cgccgtgccg acccggggcg agacgattgg gaagaagacc gcttcaagcg cctgggaccc	3780
ggcaaccgtt atgcgaatcc ttcgggaccc gcgtattgcg ggcttcgccg ctgaggtgat	3840
ctacaagaag aagccggacg gcacgccgac cacgaagatt gagggttacc gcattcagcg	3900
cgacccgatc acgctccggc cggtcgagct tgattgcgga ccgatcatcg agcccgtga	3960
gtggatatgag cttcaggcgt ggttggacgg cagggggcgc ggcaaggggc tttcccgggg	4020
gcaagccatt ctgtccgcca tggacaagct gtactgagcg tgtggcgccg tcatgacttc	4080
gaagcgccgg gaagaatcga tcaaggactc ttaccgctgc cgtcgccgga aggtggtcga	4140
cccgtccgca cctgggcagc acgaaggcac gtgcaacgtc agcatggcgg cactcgacaa	4200
gttcgttgcg gaacgcattc tcaacaagat caggcacgcc gaaggcgacg aagagacgtt	4260
ggcgcttctg tgggaagccg cccgacgctt cggcaagctc actgaggcgc ctgagaagag	4320
cggcgaacgg gcgaaccttg ttgctggagcg cgccgacgcc ctgaacgccc ttgaagagct	4380
gtacgaagac cgcgcggcag gcgcgtacga cggaccctt ggcaaggaagc acttccggaa	4440
gcaacaggca gcgctgacgc tccggcagca aggggcggaa gagcggcttg ccgaacttga	4500
agccgccgaa gccccgaagc ttccccttga ccaatggctc cccgaagacg ccgacgctga	4560
cccgaaccgg cctaagtcgt ggtggggggcg cgcgtcagta gacgacaagc gcgtgttcgt	4620
cgggctcttc gtagacaaga tcgttgtcac gaagtcgact acgggcaggg ggcaaggaa	4680
gcccacgag aagcgcgctt cgatcacgtg ggcgaagccg ccgaccgacg acgacgaaga	4740
cgacgcccag gacggcacgg aagacgtagc ggcgtagcga gacaccgga tccctcgagg	4800
ggccctattc tatagtgtca cctaaatgct agagctcgct gatcagcctc gactgtgcct	4860
tctagtgtgc agccatctgt tgtttgcccc tccccgtgc cttccttgac cctggaaggt	4920
gccactccca ctgtcctttc ctaataaaat gaggaaattg catcgattg tctgagtagg	4980
tgatcattcta ttctgggggg tgggggtggg caggacagca agggggagga ttgggaagac	5040
aatagcaggc atgctgggga tgcgggtggg tctatggctt ctgaggcgga aagaaccagg	5100
tgcccagtca tagccgaata gcctctccac ccaagcggcc ggagaacctg cgtgcaatcc	5160
actggggggcg cg	5172

<210> 10
<211> 6233

025CIP4 PCT SEQ List.txt

<212> DNA

<213> Artificial sequence

<220>

<223> Plasmid PCR-XL-TOPO-CMV-pur-attB

<400> 10

agcgcccaat acgcaaaccg cctctccccg cgcgttggcc gattcattaa tgcagctggc	60
acgacagggt tcccgactgg aaagcgggca gtgagcgcaa cgcaattaat gtgagttagc	120
tcactcatta ggcaccccgag gctttacact ttatgcttcc ggctcgtatg ttgtgtggaa	180
ttgtgagcgg ataacaattt cacacaggaa acagctatga ccatgattac gccaagctat	240
ttaggtgacg cgttagaata ctcaagctat gcatcaagct tggtagcgag ctcggatcca	300
ctagtaacgg ccgccagtggt gctggaattc gcccttggcc gcaataaaaat atctttat	360
tcattacatc tgttgtgttg ttttttgtgt gaatcgatag tactaacata cgctctccat	420
caaaacaaaa cgaaacaaaa caaactagca aaataggctg tccccagtg c aagtgcagg	480
gccagaacat ttctctatcg ataggtaccg agctcttacg cgtgctagcc ctcgagcagg	540
atctatacat tgaatcaata ttggcaatta gccatattag tcattgggta tatagcataa	600
atcaatatgg gctattggcc attgcatacg ttgtatctat atcataatat gtacatttat	660
attggctcat gtccaatatg accgccatgt tgacattgat tattgactag ttattaatag	720
taatcaatta cggggtcatt agttcatagc ccatatatgg agttccgcgt tacataactt	780
acggtaaatg gcccgctgg ctgaccgcc aacgacccc gccattgac gtcaataatg	840
acgtatgttc ccatagtaac gccaataggg actttccatt gacgtcaatg ggtggagtat	900
ttacggtaaa ctgcccactt ggcagtacat caagtgtatc atatgccaa g tccgccccct	960
attgacgtca atgacggtaa atggcccgcc tggcattatg ccagtagcat gaccttacgg	1020
gactttccta cttggcagta catctacgta ttagtcatcg ctattaccat ggtgatgcgg	1080
ttttggcagt acatcaatgg gcgtggatag cggtttgact cacggggatt tccaagtctc	1140
caccccatgg acgtcaatgg gagtttggtt tggcaccaaa atcaacggga ctttcacaaa	1200
tgctgtaaca actccgcccc attgacgcaa atgggcggta ggcgtgtacg gtgggaggtc	1260
tatataagca gagctcgttt agtgaaccgt cagatcgctt ggagacgcca tccacgctgt	1320
tttgacctcc atagaagaca ccgggaccga tccagcctcc cctcgaagct cgactctagg	1380
ggctcgagat ctgcgatcta agtaagcttg catgcctgca ggtcggccgc cacgaccggt	1440
gccgccacca tcccctgacc cacgcccctg acccctcaca aggagacgac cttccatgac	1500
cgagtacaag cccacgggtgc gcctcgccac ccgcgacgac gtcccccg g cgtacgcac	1560
cctcgccgcc gcgttcgccg actacccgcg cacgcgccac accgtcgacc cggaccgcca	1620
catcgagcgg gtcaccgagc tgcaagaact cttcctcacg cgcgtcgggc tcgacatcgg	1680

025CIP4 PCT SEQ List.txt

caaggtgtgg gtcgcggaacg acggcgccgc ggtggcggtc tggaccacgc cggagagcgt	1740
cgaagcgggg gcggtgttcg ccgagatcgg ccgcgcgatg gccgagttga gcggttccccg	1800
gctggccgcg cagcaacaga tggaaggcct cctggcgccg caccggccca aggagcccgc	1860
gtggttcctg gccaccgtcg gcgtctcgcc cgaccaccag ggcaagggtc tgggcagcgc	1920
cgctgtgctc cccggagtgg agggcgccga gcgcgcgggg gtgcccgcct tcctggagac	1980
ctccgcgccc cgcaacctcc ccttctacga gcggctcggc ttcaccgtca ccgccgacgt	2040
cgaggtgccc gaaggaccgc gcacctgggt catgaccgcg aagcccgggt cctgacgccc	2100
gccccacgac ccgcagcgcc cgaccgaaag gagcgcacga ccccatggct ccgaccgaag	2160
ccgaccgggg cggccccgcc gaccccgcac ccgccccga ggcccaccga ctctagagtc	2220
ggggcgggcg gccgcttcga gcagacatga taagatacat tgatgagttt ggacaaacca	2280
caactagaat gcagtgaata aatgcttta ttgtgaaat ttgtgatgct attgctttat	2340
ttgtaaccat tataagctgc aataaacaag ttaacaacaa caattgcatt cattttatgt	2400
ttcaggttca gggggaggtg tgggaggttt tttaaagcaa gtaaacctc taaaaatgtg	2460
gtaaaatcga taaggatcaa ttcggcttca ggtaccgtcg acgatgtagg tcacgggtctc	2520
gaagccgcgg tgcgggtgcc agggcggtgcc cttgggctcc ccgggcgcgt actccacctc	2580
acccatctgg tccatcatga tgaacgggtc gaggtggcgg tagttgatcc cggcgaacgc	2640
gcggcgcacc gggaagccct cgccctcgaa accgctgggc gcggtggtca cggtgagcac	2700
gggacgtgcg acggcgctcg cgggtgcgga tacgcggggc agcgtcagcg ggttctcgac	2760
ggtcacggcg ggcattgtcg cagccgaatt gatccgtcga ccgatgccct tgagagcctt	2820
caaccagtc agctccttcc ggtgggcgcg gggcatgact atcgctcgcc cacttatgac	2880
tgtcttcttt atcatgcaac tcgtaggaca ggtgccggca gcgctcttcc gcttcctcgc	2940
tcactgactc gctgcgctcg gtcgttcggc tgcggcgagc ggtatcagct cactcaaagg	3000
cggtaatacg gttatccaca gaatcagggg ataacgcagg aaagaacatg aagggcgaat	3060
tctgcagata tccatcacac tggcgccgc tcgagcatgc atctagaggg cccaattcgc	3120
cctatagtga gtcgtattac aattcactgg ccgctgtttt acaacgtcgt gactgggaaa	3180
accctggcgt taccctaactt aatcgcttg cagcacatcc ccctttcgcc agctggcgta	3240
atagcgaaga ggcccgcacc gatcgccctt cccaacagtt gcgcagccta tacgtacggc	3300
agtttaaggt ttacacctat aaaagagaga gccgttatcg tctgtttgtg gatgtacaga	3360
gtgatattat tgacacgccg gggcgacgga tggatgaccc cctggccagt gcacgtctgc	3420
tgtcagataa agtctcccgt gaactttacc cgggtggtgca tatcggggat gaaagctggc	3480
gcatgatgac caccgatatg gccagtgtgc cggctctccgt tatcggggaa gaagtggctg	3540
atctcagcca ccgcgaaaat gacatcaaaa acgccattaa cctgatgttc tggggaatat	3600

025CIP4 PCT SEQ List.txt

aaatgtcagg	catgagatta	tcaaaaagga	tcttcaccta	gataccttttc	acgtagaaag	3660
ccagtccgca	gaaacggtgc	tgaccccgga	tgaatgtcag	ctactgggct	atctggacaa	3720
gggaaaacgc	aagcgcaaag	agaaagcagg	tagcttgcag	tgggcttaca	tggcgatagc	3780
tagactgggc	ggtttttatg	acagcaagcg	aaccggaatt	gccagctggg	gcgccctctg	3840
gtaaggttgg	gaagccctgc	aaagtaaact	ggatggcttt	ctcgccgcca	aggatctgat	3900
ggcgagggg	atcaagctct	gatcaagaga	caggatgagg	atcgtttcgc	atgattgaac	3960
aagatggatt	gcacgcaggt	tctccggccg	cttgggtgga	gaggctattc	ggctatgact	4020
gggcacaaca	gacaatcggc	tgctctgatg	ccgccgtgtt	ccggctgtca	gcgcagggggc	4080
gcccggttct	ttttgtcaag	accgacctgt	ccggtgccct	gaatgaactg	caagacgagg	4140
cagcgcggt	atcgtaggctg	gccacgacgg	gcgttccttg	cgagctgtg	ctcgacgttg	4200
tactgaagc	gggaagggac	tggctgctat	tgggcgaagt	gccggggcag	gatctcctgt	4260
catctcacct	tgctcctgcc	gagaaagtat	ccatcatggc	tgatgcaatg	cggcggtgc	4320
atagcttga	tccggctacc	tgccattcg	accaccaagc	gaaacatcgc	atcgagcgag	4380
cacgtactcg	gatggaagcc	ggtcttgtcg	atcaggatga	tctggacgaa	gagcatcagg	4440
ggctcgcgcc	agccgaactg	ttcgccaggc	tcaaggcgag	catgcccgc	ggcgaggatc	4500
tcgtcgtgac	ccatggcgat	gcctgcttgc	cgaatatcat	ggtggaaaat	ggccgctttt	4560
ctggattcat	cgactgtggc	cggctgggtg	tggcggaccg	ctatcaggac	atagcgttgg	4620
ctacccgtga	tattgctgaa	gagcttggcg	gcgaatgggc	tgaccgcttc	ctcgtgcttt	4680
acggtatcgc	cgctcccgat	tcgcagcgca	tcgccttcta	tcgccttctt	gacgagttct	4740
tctgaattat	taacgcttac	aatttcctga	tgcggtat	tctccttacg	catctgtgcg	4800
gtatttcaca	ccgcatacag	gtggcacttt	tcggggaaat	gtgcgcggaa	cccctatttg	4860
tttatttttc	taaatacatt	caaatatgta	tccgctcatg	agacaataac	cctgataaat	4920
gcttcaataa	tagcacgtga	ggagggccac	catggccaag	ttgaccagtg	ccgttccggt	4980
gctcaccgcg	cgcgacgtcg	ccggagcgg	cgagttctgg	accgaccggc	tcgggttctc	5040
ccgggacttc	gtggaggacg	acttcgccgg	tgtggtccgg	gacgacgtga	ccctgttcat	5100
cagcgcggtc	caggaccagg	tgggtgccga	caacaccctg	gcctgggtgt	gggtgcgcgg	5160
cctggacgag	ctgtacgccg	agtggtcgga	ggtcgtgtcc	acgaacttcc	gggacgcctc	5220
cgggccggcc	atgaccgaga	tcggcgagca	gccgtggggg	cgggagttcg	ccctgcgcga	5280
ccgggccggc	aactgcgtgc	acttcgtggc	cgaggagcag	gactgacacg	tgctaaaact	5340
tcatttttaa	tttaaaagga	tctaggtgaa	gataatctca	tgaccaaact		5400
cccttaacgt	gagttttcgt	tccactgagc	gtcagacccc	gtagaaaaga	tcaaaggatc	5460

025CIP4 PCT SEQ List.txt

```

ttcttgagat ccttttttttc tgcgcgtaat ctgctgcttg caaacaacaaa aaccaccgct 5520
accagcgggtg gtttggtttgc cggatcaaga gctaccaact ctttttccga aggtaactgg 5580
cttcagcaga gcgcagatac caaatactgt ctttctagt tagccgtagt taggccacca 5640
cttcaagaac tctgtagcac cgcctacata cctcgctctg ctaatcctgt taccagtggc 5700
tgctgccagt ggcgataagt cgtgtcttac cgggttggaac tcaagacgat agttaccgga 5760
taaggcgcag cggtcgggct gaacggggggg ttcgtgcaca cagcccagct tggagcgaac 5820
gacctacacc gaactgagat acctacagcg tgagctatga gaaagcgcca cgcttcccga 5880
agggagaaaag gcggacaggt atccggtaag cggcaggggc ggaacaggag agcgcacgag 5940
ggagcttcca gggggaaacg cctgggtatct ttatagtcct gtcggggtttc gccacctctg 6000
acttgagcgt cgatttttgt gatgctcgtc agggggggcgg agcctatgga aaaacgccag 6060
caacgcggcc tttttacggt tcctgggctt ttgctggcct tttgctcaca tgttctttcc 6120
tgcgttatcc cctgattctg tggataaccg tattaccgcc tttgagttag ctgataccgc 6180
tcgccgcagc cgaacgaccg agcgcagcga gtcagttagc gaggaagcgg aag 6233

```

<210> 11
 <211> 234
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> attP containing polynucleotide

```

<400> 11
gactagtact gacggacaca ccgaagcccc ggcggaacc ctcagcggat gccccggggc 60
ttcacgtttt cccagggtcag aagcggtttt cgggagtagt gcccactg gggtaacctt 120
tgagttctct cagttggggg cgtagggctc cgcacatgac acaaggggtt gtgaccgggg 180
tggacacgta cgcgggtgct tacgaccgtc agtcgcgcga gcgcgactag taca 234

```

<210> 12
 <211> 26
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> Primer attB-for

```

<400> 12
taccgtcgac gatgtaggtc acggtc 26

```

<210> 13
 <211> 11
 <212> PRT
 <213> SV40

<400> 13

025CIP4 PCT SEQ List.txt
Cys Gly Gly Pro Lys Lys Lys Arg Lys Val Gly
1 5 10

INTERNATIONAL SEARCH REPORT

International Application No

PCT/US2005/006072

A. CLASSIFICATION OF SUBJECT MATTER
IPC 7 A01K67/027

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

IPC 7 A01K

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)

EPO-Internal, WPI Data, PAJ, BIOSIS, CHEM ABS Data, Sequence Search

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	THYAGARAJAN B ET AL: "Site-specific genomic integration in mammalian cells mediated by phage phiC31 integrase." MOLECULAR AND CELLULAR BIOLOGY. JUN 2001, vol. 21, no. 12, June 2001 (2001-06), pages 3926-3934, XP002330999 ISSN: 0270-7306 the whole document	1-142
Y	WO 03/024199 A (AVIGENICS, INC; RAPP, JEFFREY, C; SUTRAVE, PRAMOD) 27 March 2003 (2003-03-27) the whole document	1-142
Y	WO 03/025146 A (AVIGENICS, INC; RAPP, JEFFREY, C; CHRISTMANN, LEANDRO) 27 March 2003 (2003-03-27) the whole document	1-142
	----- -/-	

☒ Further documents are listed in the continuation of box C.

☒ Patent family members are listed in annex.

* Special categories of cited documents:

A document defining the general state of the art which is not considered to be of particular relevance

E earlier document but published on or after the international filing date

L document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

O document referring to an oral disclosure, use, exhibition or other means

P document published prior to the international filing date but later than the priority date claimed

T later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

X document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

Y document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.

G document member of the same patent family

Date of the actual completion of the international search

9 June 2005

Date of mailing of the international search report

06/07/2005

Name and mailing address of the ISA

European Patent Office, P.B. 5818 Patentlaan 2
NL - 2280 HV Rijswijk
Tel. (+31-70) 340-2040, Tx. 31 651 epo nl,
Fax: (+31-70) 340-3016

Authorized officer

Novak-Giese, S

INTERNATIONAL SEARCH REPORT

International Application No
PCT/US2005/006072

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	HARVEY AJ ET AL.: "Expression of exogenous protein in the egg white of transgenic chicken" NATURE BIOTECHNOLOGY, vol. 19, April 2002 (2002-04), pages 396-399, XP002973268 the whole document	1-142
A	----- WO 99/19472 A (UNIVERSITY OF GEORGIA RESEARCH FOUNDATION) 22 April 1999 (1999-04-22) the whole document	1-142
A	----- IVARIE R: "Avian transgenesis: progress towards the promise" TRENDS IN BIOTECHNOLOGY, ELSEVIER PUBLICATIONS, CAMBRIDGE, GB, vol. 21, no. 1, January 2003 (2003-01), pages 14-19, XP002983884 ISSN: 0167-7799 the whole document	
P,X	----- WO 2004/092351 A (AVIGENICS, INC) 28 October 2004 (2004-10-28) the whole document	1-142
P,X	----- WO 2004/080162 A (AVIGENICS, INC) 23 September 2004 (2004-09-23) the whole document	1-142
P,X	----- WO 2004/067713 A (AVIGENICS INC; UNIVERSITY OF GEORGIA RESEARCH FOUNDATION INC; IVARIE,) 12 August 2004 (2004-08-12) the whole document	1-142

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US2005/006072

Box II Observations where certain claims were found unsearchable (Continuation of item 2 of first sheet)

This International Search Report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:

1. ☒ Claims Nos.: **1-142 (partially)**
because they relate to subject matter not required to be searched by this Authority, namely:

The subject-matter of claims 1-142 (partially) could not be searched since it relates also to transgenic humans and the modification of the germ line of humans. The claims should be restricted to 'non-human' vertebrates.
2. ☐ Claims Nos.:
because they relate to parts of the International Application that do not comply with the prescribed requirements to such an extent that no meaningful International Search can be carried out, specifically:
3. ☐ Claims Nos.:
because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).

Box III Observations where unity of invention is lacking (Continuation of item 3 of first sheet)

This International Searching Authority found multiple inventions in this International application, as follows:

1. ☐ As all required additional search fees were timely paid by the applicant, this International Search Report covers all searchable claims.
2. ☐ As all searchable claims could be searched without effort justifying an additional fee, this Authority did not invite payment of any additional fee.
3. ☐ As only some of the required additional search fees were timely paid by the applicant, this International Search Report covers only those claims for which fees were paid, specifically claims Nos.:
4. ☐ No required additional search fees were timely paid by the applicant. Consequently, this International Search Report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.:

Remark on Protest

- ☐ The additional search fees were accompanied by the applicant's protest.
- ☐ No protest accompanied the payment of additional search fees.

INTERNATIONAL SEARCH REPORT

Information on patent family members

International Application No

PCT/US2005/006072

Patent document cited in search report		Publication date	Patent family member(s)	Publication date
WO 03024199	A	27-03-2003	WO 03024199 A2	27-03-2003
WO 03025146	A	27-03-2003	EP 1438401 A2	21-07-2004
			WO 03025146 A2	27-03-2003
			US 2003126629 A1	03-07-2003
			EP 1478751 A2	24-11-2004
			WO 02079447 A2	10-10-2002
			US 2003140363 A1	24-07-2003
WO 9919472	A	22-04-1999	AU 752946 B2	03-10-2002
			AU 1189999 A	03-05-1999
			CA 2307840 A1	22-04-1999
			EP 1023442 A1	02-08-2000
			JP 2001520009 T	30-10-2001
			WO 9919472 A1	22-04-1999
			US 6730822 B1	04-05-2004
			US 2004019922 A1	29-01-2004
			US 2004019923 A1	29-01-2004
			US 2004158882 A1	12-08-2004
WO 2004092351	A	28-10-2004	US 2004210954 A1	21-10-2004
			US 2004255345 A1	16-12-2004
			WO 2004092351 A2	28-10-2004
			US 2005034186 A1	10-02-2005
			US 2005066383 A1	24-03-2005
			WO 2004080162 A2	23-09-2004
			WO 2005040215 A2	06-05-2005
			US 2004226057 A1	11-11-2004
			US 2005003414 A1	06-01-2005
WO 2004080162	A	23-09-2004	US 2004210954 A1	21-10-2004
			US 2004255345 A1	16-12-2004
			US 2005034186 A1	10-02-2005
			US 2005066383 A1	24-03-2005
			WO 2004080162 A2	23-09-2004
			WO 2004092351 A2	28-10-2004
			WO 2005040215 A2	06-05-2005
WO 2004067713	A	12-08-2004	US 2004019922 A1	29-01-2004
			US 2004019923 A1	29-01-2004
			WO 2004067713 A2	12-08-2004